



STIC Search Report

Biotech-Chem Library

STIC Database Tracking Number: 137724

TO: Patricia Duffy
Location: rem/3b05/3c18
Art Unit: 1645
Monday, March 29, 2004

Case Serial Number: 09/438185

From: Peggy Ruppel
Location: Biotech-Chem Library
Phone: 571-272-2557
REM E01b65
peggy.ruppel@uspto.gov

Search Notes

Dear Examiner Duffy:

The results of your search request are attached.

Feel free to contact me if you have any questions.

Thank you for using STIC services.

Peggy Ruppel
2-2557

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 25, 2004, 14:10:43 ; Search time 59 Seconds
(without alignments)
1733.597 Million cell updates/sec

Title: US-09-438-185A-1047

Perfect score: 1889

Sequence: 1 VHYCERTLPKYLKIALK.....ESIPLYNQKYLSGFEVLQ 362

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_29Jan04.*

1: Geneseqp1980s.*

2: Geneseqp1990s.*

3: Geneseqp2000s.*

4: Geneseqp2001s.*

5: Geneseqp2002s.*

6: Geneseqp2003as.*

7: Geneseqp2003bs.*

8: Geneseqp2004s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|--------|-------------|--------|-------|---------------------|
| 1 | 1889 | 100.0 | 362 | 5 | ABB90626 Chlamydia |
| 2 | 1250.5 | 66.2 | 259 | 2 | AAY78593 Chlamydia |
| 3 | 237 | 12.5 | 429 | 4 | AAE11086 Protein e |
| 4 | 235.5 | 12.5 | 262 | 7 | ADD47706 Rat Prote |
| 5 | 235.5 | 12.5 | 262 | 7 | ADD48839 Rat Prote |
| 6 | 231.5 | 12.3 | 415 | 4 | AAE11161 phnAB fus |
| 7 | 231.5 | 12.3 | 444 | 4 | AAE11161 phnAB fus |
| 8 | 231.5 | 12.3 | 444 | 5 | ABG96431 Human ova |
| 9 | 231.5 | 12.3 | 466 | 5 | ABG96432 Human ova |
| 10 | 226.5 | 12.0 | 444 | 4 | AAE173208 Human pro |
| 11 | 226.5 | 12.0 | 434 | 5 | AAO17329 Neutrope s |
| 12 | 223 | 11.8 | 498 | 7 | ADE60903 Rat Prote |
| 13 | 223 | 11.8 | 498 | 7 | ADE60907 Rat Prote |
| 14 | 222 | 11.8 | 498 | 2 | AAE50185 Rat tyros |
| 15 | 221 | 11.7 | 498 | 2 | AAE36741 Rat tyros |
| 16 | 220 | 11.6 | 497 | 2 | AAE50184 Human tyr |
| 17 | 219 | 11.6 | 497 | 2 | AAE36740 Human tyr |
| 18 | 219 | 11.6 | 579 | 4 | ABE64882 Drosophil |
| 19 | 216.5 | 11.5 | 484 | 6 | AAE37558 Human hyd |
| 20 | 216.5 | 11.5 | 485 | 6 | AAE37557 Human hyd |
| 21 | 216.5 | 11.5 | 486 | 6 | AAE37556 Human hyd |
| 22 | 216.5 | 11.5 | 490 | 6 | AAE37555 Human hyd |
| 23 | 216.5 | 11.3 | 452 | 2 | AAE13119 Phenylala |
| 24 | 213.5 | 11.3 | 452 | 2 | AAW25788 Human phe |
| 25 | 213.5 | 11.3 | 452 | 2 | AAY55893 Human phe |

ALIGNMENTS

RESULT 1

ABB90626

ID ABB90626 standard; protein; 362 AA.

XX ABB90626;

XX

DT 29-AUG-2003 (revised)

DT 29-JUL-2002 (first entry)

XX

DE Chlamydia pneumoniae cp7380 protein, SEQ ID NO:201.

XX

KW Chlamydia infection; antigen; immunogen; vaccine; diagnosis;

KW human respiratory disease; cardiovascular disease; atherosclerosis;

KW coronary artery disease; carotid artery stenosis; myocardial infarction;

KW cerebrovascular disease; aortic aneurysm; claudication; stroke;

XX strain CWL029.

XX

OS Chlamydia pneumoniae.

XX

PN WO200202606-A2.

XX

PD 10-JAN-2002.

XX

PF 03-JUL-2001; 2001WO-IB001445.

XX

PR 03-JUL-2000; 2000GB-00016363.

PR

PR 11-JUL-2000; 2000GB-00017047.

PR

PR 21-JUL-2000; 2000GB-00017985.

PR

PR 07-AUG-2000; 2000GB-00019368.

PR

PR 18-AUG-2000; 2000GB-00020440.

PR

PR 14-SEP-2000; 2000GB-00022583.

PR

PR 10-NOV-2000; 2000GB-00027549.

PR

PR 22-DEC-2000; 2000GB-00031706.

XX

XX (CHIR-) CHIRON SPA.

PA

XX Ratti G, Grandi G;

PI

XX WPI; 2002-154726/20.

DR

XX N-PSDB; ABL91284.

XX

PT Novel Chlamydia pneumoniae protein useful in the manufacture of a

PT medicament for treatment or prevention of infection due to Chlamydia,

PT preferably Chlamydia pneumoniae, and for diagnostic purposes.

XX

PS Claim 1; Page 134; 364pp; English.

XX

CC Sequences ABB90526-ABB90715 represent novel proteins from Chlamydia

Aay78593 Human phe
Abb81909 Protein r
Add18674 Human dis
Aab13327 Caenorhab
Aau30703 Novel hum
Abg71947 C. violac
Abb61277 Drosophil
Abb71186 Drosophil
Aau30702 Novel hum
Aae11092 phnAB fus
Abp06168 Human ORF
Abp61194 Drosophil
Abb65685 Drosophil
Aap35686 Fungal ZB
Aab94081 Human pro
Abb59689 Drosophil
Adb65334 Human pro
Aaw69848 Amino aci
Ade54882 Rat Prote
Ade54870 Rat Prote

CC* pneumoniae (strain CMU029), and AB91184-ABL91373 represent DNA encoding
 CC them. The proteins are predicted to be immunogenic and may therefore be
 CC useful in vaccine production and for diagnostic purposes. Chlamydia
 CC pneumoniae is a common cause of respiratory disease in humans, and is
 CC also involved in the development of cardiovascular diseases such as
 CC atherosclerosis, coronary artery disease, carotid artery stenosis,
 CC myocardial infarction, cerebrovascular disease, aortic aneurysm,
 CC claudication and stroke. The proteins and nucleic acids of the invention
 CC may be used in vaccines and pharmaceutical compositions for the
 CC prevention or treatment of chlamydial infections, particularly Chlamydia
 CC pneumoniae infections. The proteins may also be used in the detection of
 CC Chlamydia pneumoniae, and the nucleic acids may be used in PCR, branched
 CC DNA probe assay or blotting techniques for determining Chlamydia
 CC pneumoniae gene expression. The present sequence represents a
 CC specifically claimed Chlamydia pneumoniae protein of the invention.
 CC (Updated on 29-AUG-2003 to standardise OS field)
 XX SQ Sequence 362 AA;

Query Match 100.0%; Score 1889; DB 5; Length 362;
 Best Local Similarity 100.0%; Pred. No. 7e-182;
 Matches 362; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VHYCERTLDPKYLKIALKLRQSLUFFQNSQLRAYSTPSYRILQKENKQALA 60
 DB 1 VHYCERTLDPKYLKIALKLRQSLUFFQNSQLRAYSTPSYRILQKENKQALA 60

QY 61 RHKISILEFFKNLLFVHLLSKNREGGCTDMVSTPFNNLWYLLSSRFSLWS 120
 DB 61 RHKISILEFFKNLLFVHLLSKNREGGCTDMVSTPFNNLWYLLSSRFSLWS 120

QY 121 YCPRFLLDYLEAFGLSDFLDHOAVIKFFPELTHFYSYYPVSGFVAPHQYLSLLQDRYFPI 180
 DB 121 YCPRFLLDYLEAFGLSDFLDHOAVIKFFPELTHFYSYYPVSGFVAPHQYLSLLQDRYFPI 180

QY 181 ASVMRTLDKDNFSLTPDLIHLLGHVPWLLHPFSSEFFINMGRLFTKVIKQVALPSKKQ 240
 DB 181 ASVMRTLDKDNFSLTPDLIHLLGHVPWLLHPFSSEFFINMGRLFTKVIKQVALPSKKQ 240

QY 241 RIOTLQSNLIAIVRCFWFTVESGLIENHEGRKAYGAVLISSPOELGHAFTDNVRVLPLEL 300
 DB 241 RIOTLQSNLIAIVRCFWFTVESGLIENHEGRKAYGAVLISSPOELGHAFTDNVRVLPLEL 300

QY 301 DQIIRLPNTSTPQETLFSIRHFDDELVELTSKLEWMLDQGLLESIPLYNQEKYLSGFEVL 360
 DB 301 DQIIRLPNTSTPQETLFSIRHFDDELVELTSKLEWMLDQGLLESIPLYNQEKYLSGFEVL 360

QY 361 CQ 362
 DB 361 CQ 362

RESULT 2
 AAY35703
 ID AAY35703 standard; protein; 259 AA.
 XX AC AAY35703;
 XX DT 17-OCT-2003 (revised)
 XX DT 13-SEP-1999 (first entry)
 XX XX
 XX Chlamydia pneumoniae transmembrane protein sequence.
 XX Respiratory disease; pneumonia; bronchitis; heart disease; sarcoidosis;
 KW sinusitis; purulent otitis media; erythema nodosum; pharyngitis; vaccine;
 KW neutralising epitope.
 XX Chlamydophila pneumoniae.
 OS OS
 XX WO9927105-A2.
 PN PN
 XX 03-JUN-1999.
 PD PD
 XX XX

PF 20-NOV-1998; 98WO-IB001890.
 XX 21-NOV-1997; 97FR-00014673.
 PR 04-NOV-1998; 98US-0107078P.
 XX (GEST) GENSET.
 XX Griffais R;
 FI WPI; 1999-357842/30.
 DR Genome sequence of Chlamydia pneumoniae.
 XX Page 1410-1411; Disclosure; 1913pp; English.
 CC AAY34584-Y35879 represent the proteins encoded by all the open reading
 CC frames in the complete genome (see AAX91990) of Chlamydia pneumoniae. C.
 CC pneumoniae causes respiratory disease such as pneumonia and bronchitis,
 CC and is thought to be a contributing factor in heart disease, sarcoidosis,
 CC sinusitis, purulent otitis media, erythema nodosum or pharyngitis. The
 CC polypeptides encoded by the open reading frames of the C. pneumoniae
 CC genome (see AAY34584-Y35879) can be used in immunogenic compositions as
 CC vaccines. Vectors containing C. pneumoniae nucleotide sequences can also
 CC be used as immunogenic compositions, especially where the vector directs
 CC the expression of a neutralising epitope of C. pneumoniae. (Updated on 17
 CC -OCT-2003 to standardise OS field)
 XX SQ Sequence 259 AA;

Query Match 66.2%; Score 1250.5; DB 2; Length 259;
 Best Local Similarity 91.8%; Pred. No. 1.4e-117;
 Matches 245; Conservative 3; Mismatches 8; Indels 11; Gaps 1;

QY 96 VVSTPFFNRNLWYLLSSRFSLWSKSYCPREFLDYLEAFGLSDFLDHOAVIKFFPELTHF 155
 DB 4 MVSFPFLTVFSMEKLLS-----KIFLDYLEAFGLSDFLDHOAVIKFFPELTHF 52

QY 156 SYYPVSGFVAPHQYLSLLQDRYFPIASVMRTLDKDNFSLTPDLIHLLGHVPWLLHPFSFS 215
 DB 53 SYYPVSGFVAPHQYLSLLQDRYFPIASVMRTLDKDNFSLTPDLIHLLGHVPWLLHPFSFS 112

QY 216 EFFINMGRLFTKVIKQVALPSKKORIQTLOSNIAlAVRCFWFTVESGLIENHEGRKAYG 275
 DB 113 EFFINMGRLFTKVIKQVALPSKKORIQTLOSNIAlAVRCFWFTVESGLIENHEGRKAYG 172

QY 276 AVLSSPOELGHAFTDNVRVLPLELDQIIRLPNTSTPQETLFSIRHFDDELVELTSKLEW 335
 DB 173 AVLSSPOELGHAFTDNVRVLPLELDQIIRLPNTSTPQETLFSIRHFDDELVELTSKLEW 232

QY 336 MLDQGLLESIPLYNQEKYLSGFEVLCO 362
 DB 233 MLDQGLLESIPLYNQEKYLSGFEVLCO 259

RESULT 3
 AAE11086
 ID AAE11086 standard; protein; 429 AA.
 XX AC AAE11086;
 XX DT 18-DEC-2001 (first entry)
 XX DE Protein encoded by pFUSAB vector DNA insert sequence.
 XX Phenylalanine hydroxylase; PAH; phhA; 4a-carbinolamine dehydratase; phhB;
 KW aromatic aminotransferase; phhC; milk protein; animal protein; casein;
 KW proteinaceous food product; globulin; whey protein; phenylketonuria; PKU;
 KW inherited metabolic disorder; impaired brain function; nootropic;
 KW cell therapy.
 XX Escherichia coli.
 OS Unidentified.
 OS Chimeric.

XX Key Location/Qualifiers
FH 1. .26
FT /label= LacZ_peptide
FT 27. .28
FT /label= Linker_peptide
FT 29. .289
FT /label= phhA protein
FT 290. .311
FT /label= Junction_peptide
FT 312. .429
FT /label= phhB protein
XX WO200168822-A2.
XX
XX 20-SEP-2001.
XX
XX 14-MAR-2001; 2001WO-DK000172.
XX
XX 14-MAR-2000; 2000US-00525116.
XX (NTLA-) NTLAB APS.
XX Johnsen M, Ravn P, Madsen SM, Vrang A, Israelsen H, Bredmose L;
PI Arnau J, Jensen SH, Gjetting T, Nielsen E;
XX
XX WPI; 2001-590055/66.
XX N-PSDB; AAD18533.
XX
XX Novel recombinant cells comprising a nucleic acid encoding a gene product
FT having phenylalanine hydroxylase activity, that is derived from a
FT prokaryotic organism, is useful for treating phenylketonuria in mammals.
XX
XX Example 2; Fig 5; 91pp; English.
XX
XX The patent discloses novel cells comprising a nucleic acid encoding a
CC gene product having phenylalanine hydroxylase (PAH) activity such as
CC phenylalanine hydroxylase (pHHA), 4a-carbinolamine dehydratase (pHNB) and
CC aromatic aminotransferase (pHNC), which are derived from a prokaryotic
CC organism. The patent also relates to fusion proteins comprising a protein
CC enhancing and/or stabilising the PAH activity in addition to PAH
CC activity. The cells are useful for producing PAH. The sequences of the
CC invention are also useful for preparing a proteinaceous food product
CC (animal protein) such as a milk protein derived from casein, globulin or a
CC whey protein) having reduced content of phenylalanine. The method
CC involves contacting the food product starting material with the cells or
CC fusion proteins such that at least part of the phenylalanine content of
CC the starting material is converted into compounds that do not cause
CC phenylketonuria (PKU) by the enzymatically active product. PAH enzyme is
CC useful for manufacturing a medicament for treating PKU, which is an
CC inherited metabolic disorder resulting in an accumulation in the body of
CC L-phenylalanine and metabolites that can cause impaired brain function.
CC The present sequence is a protein encoded by pFUSAB vector DNA insert
CC sequence. This sequence comprises a lacZ peptide from Escherichia coli, a
CC linker peptide and phhA and phhB proteins
XX
XX Sequence 429 AA;
SQ
Query Match 12.5%; Score 237; DB 4; Length 429;
Best Local Similarity 23.9%; Pred. No. 7e-15;
Matches 66; Conservative 56; Mismatches 114; Indels 40; Gaps 7;
QY 66 STLEFFKNLLFVHLLSKNREGCGSTDMVSTPFFNRNLWYLLSRRLKWSYCPRF 125
DB 20 SVLEFGFKLTQYVARPD-----DNGTHYPETEHQVNWNLITPQLKVIQRACQE 72
QY 126 FDIYLEAFGLSDFLDHQAVIKFFLE-----THFSYYPVSGFVAPHQYLSLLQDRYFP 179
DB 73 YLDGIEQLG-----LPHERIPOLDINRVLOATTGWRVARVPALIPFOTFELLASQOFP 127
QY 180 IASVVRTLDKNFSITPDLIHDLLGHVPWLLHPSSEFFINMGRLFTKVIKQVALPSKK 239
DB 128 VATFIRTPLELDLQEPDIFHEIFGHCPLTNPNWPAETHFYKGLGKA-----SKE 179

QY 240 GRIQTLQSNLLAIIVRCFWFTVESGLIENHGRKAYGAVLISSPQELGHAFID---NVRVL 296
DB 180 ER-----VFLRLYWMTFIEGLVETDQKRIYGGILSPKRTVYSLSDPELHQAFN 231
QY 297 PLELDQIIRLPENTSTPQETLFSIRHFDLVELVTSK 332
DB 232 PLE---AMRTPYRIDILQPLVFLPDLKRLFLAQE 264
RESULT 4
ADD47706
ID ADD47706 standard; protein; 262 AA.
XX
XX ADD47706;
XX
XX 29-JAN-2004 (first entry)
XX Rat Protein 1TOH, SEQ ID NO 13402.
XX Rat; pain; neuronal tissue; Gene therapy; spinal segmental nerve injury;
KW Chronic constriction injury; CCI; spared nerve injury; SNI; Chung.
XX
XX Rattus norvegicus.
XX
XX WO2003016475-A2.
XX
XX 27-FEB-2003.
XX
XX 14-AUG-2002; 2002WO-US025765.
XX
XX 14-AUG-2001; 2001US-0312147P.
XX
XX 01-NOV-2001; 2001US-0346382P.
XX
XX 26-NOV-2001; 2001US-0333347P.
XX
XX (GCHO) GEN HOSPITAL CORP.
XX (FARB) BAYER AG.
XX Woolf C, D'urso D, Befort K, Costigan M;
XX WPI; 2003-268312/26.
XX
XX New composition comprising two or more isolated polypeptides, useful for
PT preparing a medicament for treating pain in an animal.
XX
XX Claim 1; Page; 1017pp; English.
XX
XX The invention discloses a composition comprising two or more isolated rat
CC or human polynucleotides or a polynucleotide which represents a fragment,
CC derivative or allelic variation of the nucleic acid sequence. Also
CC claimed are a vector comprising the novel polynucleotide, a host cell
CC comprising the vector, a method for identifying a nucleotide sequence
CC which is differentially regulated in an animal subjected to pain and a
CC kit to perform the method, an array, a method for identifying an agent
CC that increases or decreases the expression of the polynucleotide sequence
CC that is differentially expressed in neuronal tissue of a first animal
CC subjected to pain, a method for identifying a compound which regulates
CC the expression of a polynucleotide sequence which is differentially
CC expressed in an animal subjected to pain, a method for identifying a
CC compound that regulates the activity of one or more of the
CC polynucleotides, a method for producing a pharmaceutical composition, a
CC method for identifying a compound or small molecule that regulates the
CC activity in an animal of one or more of the polypeptides given in the
CC specification, a method for identifying a compound useful in treating
CC pain and a pharmaceutical composition comprising the one or more
CC polypeptides or their antibodies. The polynucleotide or the compound that
CC modulates its activity is useful for preparing a medicament for treating
CC pain (e.g. spinal segmental nerve injury (Chung), chronic constriction
CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene
CC therapy). The sequence presented is a rat protein (shown in Table 2 of
CC the specification) which is differentially expressed during pain. Note:
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic form directly from WIPO at

CC- ftp.wipo.int/pub/published_pct_sequences.

XX SQ Sequence 262 AA;
 Query Match 12.5%; Score 235.5; DB 7; Length 262;
 Best Local Similarity 24.5%; Pred. No. 4.9e-15;
 Matches 61; Conservative 52; Mismatches 103; Indels 33; Gaps 6;
 QY 93 DMAVSTPFNNRWLWYLLSRFSLWKSQYCPFFLDYLEAFGLSLDFLDHQAVIKFFELE 152
 DB 13 DNGFIHYPETEHQVWNTLITRQLKVIIEGRACQYLDGIEQLG-----LPHRIQPDLEIN 67
 QY 153 -----THFSYYPVSGFVAPHQYLSLQDRYFPIASVMRTLDKDNFSLDFLDHLLGHV 206
 DB 68 RVLQATTGWRVARVPALIPFQTFFELLASQOFPVATFIRPELDYLOQDFIHFHGC 127
 QY 207 PWLLHPSFSEFFINMGLFTKVIKQALPSKKORIQTLOSNIATVRCFWFTVESGLIE 266
 DB 128 PLLTNPFABETHYKGLKA-----SKEER-----VFLARLYWMTIEFGLVE 171
 QY 267 NHEGRKAYGAVLISSPQELGHAFID---NVRVLPLELDQIIRLPNTSTPQETLSIRHF 323
 DB 172 TDQGRYVGGILSSPKETVYSLSDPLHQAHPLE---AMRTPYRIDILOPLVFLVDPDL 228
 QY 324 DELVELTSK 332
 DB 229 KSLFQLAQE 237

RESULT 5

ADD48839
 ID ADD48839 standard; protein; 262 AA.

AC ADD48839;
 DT 29-JAN-2004 (first entry)

DE Rat Protein 1TOH, SEQ ID NO 14549.
 KW Rat; pain; neuronal tissue; gene therapy; spinal segmental nerve injury;
 KW chronic constriction injury; CCI; spared nerve injury; SNI; Chung.
 XX Rattus norvegicus.

XX WO2003016475-A2.

XX 27-FEB-2003.

XX 14-AUG-2002; 2002WO-US025765.

XX 14-AUG-2001; 2001US-0312147P.

XX 01-NOV-2001; 2001US-0346382P.

XX 26-NOV-2001; 2001US-0333347P.

XX (GEMO) GEN HOSPITAL CORP.

XX (FARB) BAYER AG.

XX Woolf C, D'urso D, Befort K, Costigan M;

XX WPI; 2003-268312/26.

XX New composition comprising two or more isolated polypeptides, useful for
 preparing a medicament for treating pain in an animal.

XX Claim 1; Page; 1017pp; English.

XX The invention discloses a composition comprising two or more isolated rat
 or human polynucleotides or a polynucleotide which represents a fragment,
 CC derivative or allelic variation of the nucleic acid sequence. Also
 CC claimed are a vector comprising the novel polynucleotide, a host cell
 CC comprising the vector, a method for identifying a nucleotide sequence
 CC which is differentially regulated in an animal subjected to pain and a
 CC kit to perform the method, an array, a method for identifying an agent

CC that increases or decreases the expression of the polynucleotide sequence
 CC that is differentially expressed in neuronal tissue of a first animal
 CC subjected to pain, a method for identifying a compound which regulates
 CC the expression of a polynucleotide sequence which is differentially
 CC expressed in an animal subjected to pain, a method for identifying a
 CC compound that regulates the activity of one or more of the
 CC polynucleotides, a method for producing a pharmaceutical composition, a
 CC method for identifying a compound or small molecule that regulates the
 CC activity in an animal of one or more of the polypeptides given in the
 CC specification, a method for identifying a compound useful in treating
 CC pain and a pharmaceutical composition comprising the one or more
 CC polypeptides or their antibodies. The polynucleotide or the compound that
 CC modulates its activity is useful for preparing a medicament for treating
 CC pain (e.g. spinal segmental nerve injury (Chung), chronic constriction
 CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene
 CC therapy). The sequence presented is a rat protein (shown in Table 2 of
 CC the specification) which is differentially expressed during pain. Note:
 CC The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic form directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 262 AA;

Query Match 12.5%; Score 235.5; DB 7; Length 262;

Best Local Similarity 24.5%; Pred. No. 4.9e-15;
 Matches 61; Conservative 52; Mismatches 103; Indels 33; Gaps 6;

QY 93 DMAVSTPFNNRWLWYLLSRFSLWKSQYCPFFLDYLEAFGLSLDFLDHQAVIKFFELE 152

DB 13 DNGFIHYPETEHQVWNTLITRQLKVIIEGRACQYLDGIEQLG-----LPHRIQPDLEIN 67

QY 153 -----THFSYYPVSGFVAPHQYLSLQDRYFPIASVMRTLDKDNFSLDFLDHLLGHV 206

DB 68 RVLQATTGWRVARVPALIPFQTFFELLASQOFPVATFIRPELDYLOQDFIHFHGC 127

QY 207 PWLLHPSFSEFFINMGLFTKVIKQALPSKKORIQTLOSNIATVRCFWFTVESGLIE 266

DB 128 PLLTNPFABETHYKGLKA-----SKEER-----VFLARLYWMTIEFGLVE 171

QY 267 NHEGRKAYGAVLISSPQELGHAFID---NVRVLPLELDQIIRLPNTSTPQETLSIRHF 323

DB 172 TDQGRYVGGILSSPKETVYSLSDPLHQAHPLE---AMRTPYRIDILOPLVFLVDPDL 228

QY 324 DELVELTSK 332

DB 229 KSLFQLAQE 237

RESULT 6

AAE11161

ID AAE11161 standard; protein; 415 AA.

XX AAE11161;

AC AAE11161;

XX 18-DEC-2001 (first entry)

XX phhAB fusion protein fragment.

XX Phenylalanine hydroxylase; PAH; phhA; 4a-carbinolamine dehydratase; phhB;
 XX aromatic aminotransferase; phhC; milk protein; animal protein; casein;
 XX proteinaceous food product; globulin; whey protein; phenylketonuria; PKU;
 XX inherited metabolic disorder; impaired brain function; nontropic;
 XX cell therapy; phhAB fusion protein.

XX Unidentified.

XX Key Location/Qualifiers

XX Region /note= "phhAB fusion peptide #1"

XX Region 55..90

XX Region /note= "phhAB fusion peptide #2"

XX Region 94..157

XX Region /note= "phhAB fusion peptide #3"

FT Region 168..223 /note="phHAB fusion peptide #4"
 FT Region 228..272 /note="phHAB fusion peptide #5"
 FT Region 295..336 /note="phHAB fusion peptide #6"
 FT Region 353..380 /note="phHAB fusion peptide #7"
 FT Region 396..415 /note="phHAB fusion peptide #8"
 XX
 XX WO200168822-A2.
 XX
 XX 20-SEP-2001.
 XX
 XX 14-MAR-2001; 2001WO-DK000172.
 XX
 XX 14-MAR-2000; 2000US-00525116.
 XX
 XX (NILA-) NILAB APS.
 XX
 XX Johnsen M, Ravn P, Madsen SM, Vrang A, Israelsen H, Bredmose L;
 XX Arnau J, Jensen SH, Gjetting T, Nielsen E;
 XX
 XX WPI; 2001-590055/66..
 XX
 XX Novel recombinant cells comprising a nucleic acid encoding a gene product
 XX having phenylalanine hydroxylase activity, that is derived from a
 XX prokaryotic organism, is useful for treating phenylketonuria in mammals.
 XX
 XX Example 8; Fig 18; 91pp; English.
 XX
 XX The patent discloses novel cells comprising a nucleic acid encoding a
 XX gene product having phenylalanine hydroxylase (PAH) activity such as
 XX phenylalanine hydroxylase (phhA), 4a-carbinolamine dehydratase (phhB) and
 XX aromatic aminotransferase (phhC), which are derived from a prokaryotic
 XX organism. The patent also relates to fusion proteins comprising a protein
 XX enhancing and/or stabilising the PAH activity in addition to PAH
 XX activity. The cells are useful for producing PAH. The sequences of the
 XX invention are also useful for preparing a proteinaceous food product
 XX (animal protein) such as a milk protein derived from casein, globulin or a
 XX whey protein) having reduced content of phenylalanine. The method
 XX involves contacting the food product starting material with the cells or
 XX fusion proteins such that at least part of the phenylalanine content of
 XX the starting material is converted into compounds that do not cause
 XX phenylketonuria (PKU) by the enzymatically active product. PAH enzyme is
 XX useful for manufacturing a medicament for treating PKU, which is an
 XX inherited metabolic disorder resulting in an accumulation in the body of
 XX L-phenylalanine and metabolites that can cause impaired brain function.
 XX The present sequence is phHAB fusion protein fragment
 XX
 XX Sequence 415 AA;
 XX
 XX Query Match 12.5%; Score 235.5; DB 4; Length 415;
 XX Best Local Similarity 24.5%; Pred. No. 9.5e-15;
 XX Matches 61; Conservative 52; Mismatches 103; Indels 33; Gaps 6;
 QY 93 DNAAVSTPPFNRLNRYLLSSRFSWYKSCPRFFLDYLAEGLLSDFLDHQAVIKFFPLE 152
 DB 26 DRGFTHYPETERQVNTLTITQLKVIQGRACQEYLDGIEQLG-----LPHERIPQLDEIN 80
 QY 153 -----THPSYYPVSGFVAPHOYLSLQDRVYPIASVNRITDKNFSITPDLIHLLGHV 206
 DB 81 RVLQATTGRVARVVALIPFQIFFELLASQFPVATFRTPELDYLOQEPDIFHFHGC 140
 QY 207 PMLHPSFSEPFINGRLTKVIEKVOALPSKQRIQLQSNLIAIVRCFTVFTVESGLIE 266
 DB 141 PLITNFWFAEFTHTYKGLGKA-----SKEER-----VFRLYNTWTFISFGIVE 184
 QY 267 NHEGRKAYGAVLISSPOELGHAFID---NVRVLPLELDQIIRLPNTSTPQETLSIRHF 323
 DB 185 TDGKRIYGGILSSPSKEIVYLSDEPLHQAENPLE---AMETPRVILQPLVFLPDL 241

QY 324 DELVELTSK 332
 DB 242 KRLFOLAQE 250
 RESULT 7
 AAM80192
 ID AAM80192 standard; protein; 444 AA.
 XX
 XX AAM80192;
 XX
 XX 06-NOV-2001 (first entry)
 XX
 XX Human protein SEQ ID NO 3838.
 XX
 XX Human; cytokine; cell proliferation; cell differentiation; gene therapy;
 XX vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
 XX tissue growth factor; immunomodulatory; cancer; leukaemia;
 XX nervous system disorder; arthritis; inflammation.
 XX
 XX Homo sapiens.
 XX
 XX WO200157190-A2.
 XX
 XX 09-AUG-2001.
 XX
 XX 05-FEB-2001; 2001WO-US004098.
 XX
 XX 03-FEB-2000; 2000US-00496914.
 XX
 XX 27-APR-2000; 2000US-00560875.
 XX
 XX 20-JUN-2000; 2000US-00598075.
 XX
 XX 19-JUL-2000; 2000US-00620325.
 XX
 XX 01-SEP-2000; 2000US-00654936.
 XX
 XX 15-SEP-2000; 2000US-0063561.
 XX
 XX 20-OCT-2000; 2000US-00693325.
 XX
 XX 30-NOV-2000; 2000US-00728422.
 XX
 XX (HYSB-) HYSEQ INC.
 XX
 XX Tang YT, Liu C, Drmanac RT, Asundi V, Zhou P, Xu C, Cao Y;
 XX Ma Y, Zhao QA, Wang D, Wang J, Zhang J, Ren F, Chen R, Wang ZW;
 XX Xue AJ, Yang Y, Wejhrman T, Goodrich R;
 XX
 XX WPI; 2001-476283/51.
 XX
 XX N-PSDB; AAK53325.
 XX
 XX Nucleic acids encoding polypeptides with cytokine-like activities, useful
 XX in diagnosis and gene therapy.
 XX
 XX Claim 20; Page 446-447; 6221pp; English.
 XX
 XX The invention relates to polynucleotides (AAK51456-AAK53435) and the
 XX encoded polypeptides (AAM78323-AAK80302) that exhibit activity elating to
 XX cytokine, cell proliferation or cell differentiation or which may induce
 XX production of other cytokines in other cell populations. The
 XX polynucleotides and polypeptides are useful in gene therapy, vaccines or
 XX peptide therapy. The polypeptides have various cytokine-like activities,
 XX e.g. stem cell growth factor activity, haematopoiesis regulating
 XX activity, tissue growth factor activity, immunomodulatory activity and
 XX activin/inhibin activity and may be useful in the diagnosis and/or
 XX treatment of cancer, leukaemia, nervous system disorders, arthritis and
 XX inflammation. Note: Records for SEQ ID NO 2110 (AAK52581), 2111
 XX (AAK52582) and 3666 (AAM80020) are omitted as the relevant pages from the
 XX sequence listing were missing at the time of publication
 XX
 XX Sequence 444 AA;
 XX
 XX Query Match 12.3%; Score 231.5; DB 4; Length 444;
 XX Best Local Similarity 27.8%; Pred. No. 2.6e-14;
 XX Matches 64; Conservative 43; Mismatches 96; Indels 27; Gaps 4;
 QY 107 WYRLSSRFSLSWYKSCPRFFLDYLAEGLLSDFLDH-----QAVIKFELETSPSY 158

PR 26-SEP-2001; 2001US-0325102P.
PR 26-SEP-2001; 2001US-0325149P.
XX (MILL-) MILLENNIUM PHARM INC.
PA Monahan JE, Gannavarapu M, Hoersch S, Kamatkar S, Kovatis SG;
PI Meyers RE, Morrissey MP, Olandt PJ, Sen A, Vieby PO, Mills GB;
XX Bat RC, Lu K, Schmandt RE, Zhao X, Glatt K;
XX WPI; 2002-723277/78.
DR N-PSDB; ABS76531.
XX Assessing whether a patient is afflicted with ovarian cancer, useful in
XX assessing the stage or progression of the disease, comprises comparing
XX the expression level of a cancer marker in a sample from a patient and
XX from a non cancer patient.
XX Disclosure; Page 443-444; 481pp; English.
XX
XX The present invention relates to a new method for assessing whether a
XX patient is afflicted with ovarian cancer. The method involves comparing
XX the expression level of a marker in a patient sample and the normal level
XX of expression of the marker in a control non-ovarian cancer sample, where
XX the marker is selected from 363 cancer markers described in the
XX specification. The method of the invention is useful in diagnosing or
XX characterising cancer, in detecting the presence of cancer as early as
XX possible, and the recurrence of ovarian cancer. The method may also be of
XX particular use with patients having an enhanced risk of developing
XX ovarian cancer (e.g. patients having a familial history of ovarian
XX cancer). The cancer markers may be used in the management and treatment
XX of e.g. brain and central nervous system disorders (e.g. bacterial and
XX viral meningitis, Alzheimer's disease or Parkinson's disease), brain
XX disorders (e.g. cerebral oedema, hydrocephalus or brain herniations),
XX inflammations (e.g. bacterial or viral meningitis or encephalitis),
XX testicular disorders (e.g. nontuberculous granulomatous orchitis),
XX connective tissue disorders, or heart disorders (e.g. ischaemic heart
XX disease or atherosclerosis). The compositions and methods may also be
XX used in assessing the histological type of neoplasm associated with
XX ovarian cancer, monitoring the progression of ovarian cancer, determining
XX whether ovarian cancer has metastasized or is likely to metastasize,
XX selecting a composition for inhibiting ovarian cancer, assessing the
XX ovarian carcinogenic potential of a compound, or inhibiting ovarian
XX cancer or at risk of developing ovarian cancer. The present amino acid
XX sequence represents one of the ovarian cancer markers described in the
XX invention
XX
XX Sequence 466 AA;
Query Match 12.3%; Score 231.5; DB 5; Length 466;
Best Local Similarity 27.8%; Pred. No. 2.8e-14;
Matches 64; Conservative 43; Mismatches 96; Indels 27; Gaps 4;
QY 107 WYRLSSRFLSKWYCPFFLDYLEAFGLSDFLDH-----QAVIKFFELTHFSY 158
DB 174 WGVFQELNKLXPTACR---EYLKXNPLLSKYCGYREDNIPQLEDVSNFLKERTGSIR 230
QY 159 PVSGFVAHQYLSLLQDRYFPFIASVWRITLDKDNFSLTPDLHDLGHVPMLLHPSSEFF 218
DB 231 PVAGYLSRDLFSLGAFVFRHCTQVYRHSDFYTPEDTCHELLGHVPLLAEPFQAQS 290
QY 219 INNGRLFTKVIKVOALPSKQRIQTLQSNLIAIVRCFWFTVESGLIENHEGRKAYGAVL 278
DB 291 QEIG-----LASLGASEAVQKLT-----CYFFTFEGLCKDQQLRVFGAGL 334
QY 279 ISSPOLGHAFIDNVRVPLELDQIIRLPNTSTPQETLFIIRHDELVE 328
DB 335 LSSISLKHGALSHGAKVPFDPKTKCKOECUITTFQDVYVSESFEDEAKE 384
RESULT 10
ID AAM79208
ID AAM79208 standard; protein; 444 AA.
XX

AC AAM79208;
XX 06-NOV-2001 (first entry)
DE Human protein SEQ ID NO 1870.
XX Human; cytokine; cell proliferation; cell differentiation; gene therapy;
KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
KW tissue growth factor; immunomodulatory; cancer; leukaemia;
KW nervous system disorder; arthritis; inflammation.
XX Homo sapiens.
OS
XX WO200157190-A2.
PN
XX 09-AUG-2001.
PD
XX 05-FEB-2001; 2001WO-US004098.
PF
XX 03-FEB-2000; 2000US-00496914.
PR
XX 27-APR-2000; 2000US-00560875.
PR
XX 20-JUN-2000; 2000US-00598075.
PR
XX 19-JUL-2000; 2000US-00620325.
PR
XX 01-SEP-2000; 2000US-00654936.
PR
XX 15-SEP-2000; 2000US-00663581.
PR
XX 30-OCT-2000; 2000US-00693325.
PR
XX 30-NOV-2000; 2000US-00728422.
XX (HYSE-) HYSEQ INC.
PA
XX Tang YT, Liu C, Drmanac RT, Asundi V, Zhou P, Xu C, Cao Y;
PI Ma Y, Zhao QA, Wang D, Wang J, Zhang J, Ren F, Chen R, Wang ZW;
PI Xue AJ, Yang Y, Wejhrman T, Goodrich R;
XX WPI; 2001-476283/51.
DR N-PSDB; AAK52341.
XX
XX Nucleic acids encoding polypeptides with cytokine-like activities, useful
XX in diagnosis and gene therapy.
XX
XX Claim 20; Page 4255-4256; 6221pp; English.
XX
XX The invention relates to polynucleotides (AAK51456-AAK53435) and the
XX encoded polypeptides (AAM78323-AAM80302) that exhibit activity elating to
XX cytokine, cell proliferation or cell differentiation or which may induce
XX production of other cytokines in other cell populations. The
XX polynucleotides and polypeptides are useful in gene therapy, vaccines or
XX peptide therapy. The polypeptides have various cytokine-like activities,
XX e.g. stem cell growth factor activity, haematopoiesis regulating
XX activity, tissue growth factor activity, immunomodulatory activity and
XX activin/inhibin activity and may be useful in the diagnosis and/or
XX treatment of cancer, leukaemia, nervous system disorders, arthritis and
XX inflammation. Note: Records for SEQ ID NO 2110 (AAK52581), 2111
XX (AAK52582) and 3666 (AAM80020) are omitted as the relevant pages from the
XX sequence listing were missing at the time of publication
XX
XX Sequence 444 AA;
Query Match 12.0%; Score 227.5; DB 4; Length 444;
Best Local Similarity 27.4%; Pred. No. 6.7e-14;
Matches 63; Conservative 43; Mismatches 97; Indels 27; Gaps 4;
QY 107 WYRLSSRFLSKWYCPFFLDYLEAFGLSDFLDH-----QAVIKFFELTHFSY 158
DB 174 WGVFQELNKLXPTACR---EYLKXNPLLSKYCGYREDNIPQLEDVSNFLKERTGSIR 230
QY 159 PVSGFVAHQYLSLLQDRYFPFIASVWRITLDKDNFSLTPDLHDLGHVPMLLHPSSEFF 218
DB 231 PVAGYLSRDLFSLGAFVFRHCTQVYRHSDFYTPEDTCHELLGHVPLLAEPFQAQS 290
QY 219 INNGRLFTKVIKVOALPSKQRIQTLQSNLIAIVRCFWFTVESGLIENHEGRKAYGAVL 278
DB 291 QEIG-----LASLGASEAVQKLT-----CYFFTFEGLCKDQQLRVFGAGL 334

Best Local Similarity 26.7%; Pred. No. 2.3e-13;
Matches 65; Conservative 39; Mismatches 103; Indels 36; Gaps 5;
QY 112 SSRFSLWKS-----YCPREFDYLEAFGLSDFLDH-----QAVIKFEELETHF 155
DB 227 ABEIATWKEVYVTLKGLYATHACREHLEGFOLLERYCYGREDSPQLEDVSRFLKERTGF 286
QY 156 SYYPVSGFVAPHQYLSLLQDRYFPIASVMRTLDKDNFSLTPDLHDLGHVPMLLHPSFS 215
DB 287 QLRPVAGLLSARDFLASLAFVFOCTQYIRHASSPMHSPEDCCHELLGHVPMADRTFA 346
QY 216 EFFINMGRLFTKVIEKQVALPSKKQRIQTLQSNLIAIVRCFWFTVESGLIENHEGRKAYG 275
DB 347 QFSQDIG-----LASLGASDEEIEKLS-----VYVFTVEFGCKQNGELKAYG 390
QY 276 AVLTSSPQELGHAFIDNVRLVPLELDQIIRLPFNSTPQETLFSIRHF-----DELVELTS 331
DB 391 AGLSSYGGELLHSLSEEPVAFDPDPTAAVQYQDTYQPVYFVSESFNDKDKLRNYAS 450
QY 332 KLE 334
DB 451 RIQ 453

RESULT 13
ADE60907
ID ADE60907 standard; protein; 498 AA.
AC ADE60907;
DT 29-JAN-2004 (first entry)
DE Rat Protein P04177, SEQ ID NO 6821.
KW Rat; pain; neuronal tissue; gene therapy; spinal segmental nerve injury;
KW chronic constriction injury; CCI; spared nerve injury; SNI; Chung.
XX Rattus norvegicus.
OS WO2003016475-A2.
PN 27-FEB-2003.
PD 14-AUG-2002; 2002WO-US025765.
PF 14-AUG-2001; 2001US-0312147P.
PR 01-NOV-2001; 2001US-0346382P.
PR 26-NOV-2001; 2001US-0333347P.
XX (GHEO) GEN HOSPITAL CORP.
PA (FARB) BAYER AG.
XX Woolf C, D'urso D, Befort K, Costigan M;
PI WPI; 2003-268312/26.
DR GENBANK; P04177.
XX New composition comprising two or more isolated polypeptides, useful for
PT preparing a medicament for treating pain in an animal.
XX
XX Claim 1; Page; 1017pp; English.

CC The invention discloses a composition comprising two or more isolated rat
CC or human polynucleotides or a polynucleotide which represents a fragment,
CC derivative or allelic variation of the nucleic acid sequence. Also
CC claimed are a vector comprising the novel polynucleotide, a host cell
CC comprising the vector, a method for identifying a nucleotide sequence
CC which is differentially regulated in an animal subjected to pain and a
CC kit to perform the method, an array, a method for identifying an agent
CC that increases or decreases the expression of the polynucleotide sequence
CC that is differentially expressed in neuronal tissue of a first animal
CC subjected to pain, a method for identifying a compound which regulates
CC the expression of a polynucleotide sequence which is differentially

CC expressed in an animal subjected to pain, a method for identifying a
CC compound that regulates the activity of one or more of the
CC polynucleotides, a method for producing a pharmaceutical composition, a
CC method for identifying a compound or small molecule that regulates the
CC activity in an animal of one or more of the polypeptides given in the
CC specification, a method for identifying a compound useful in treating
CC pain and a pharmaceutical composition comprising the one or more
CC polypeptides or their antibodies. The polynucleotide or the compound that
CC modulates its activity is useful for preparing a medicament for treating
CC pain (e.g. spinal segmental nerve injury (SNI)) in an animal (e.g. gene
CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene
CC therapy). The sequence presented is a rat protein (shown in Table 2 of
CC the specification) which is differentially expressed during pain. Note:
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic form directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 498 AA;

Query Match 11.8%; Score 223; DB 7; Length 498;
Best Local Similarity 26.7%; Pred. No. 2.3e-13;
Matches 65; Conservative 39; Mismatches 103; Indels 36; Gaps 5;
QY 112 SSRFSLWKS-----YCPREFDYLEAFGLSDFLDH-----QAVIKFEELETHF 155
DB 227 ABEIATWKEVYVTLKGLYATHACREHLEGFOLLERYCYGREDSPQLEDVSRFLKERTGF 286
QY 156 SYYPVSGFVAPHQYLSLLQDRYFPIASVMRTLDKDNFSLTPDLHDLGHVPMLLHPSFS 215
DB 287 QLRPVAGLLSARDFLASLAFVFOCTQYIRHASSPMHSPEDCCHELLGHVPMADRTFA 346
QY 216 EFFINMGRLFTKVIEKQVALPSKKQRIQTLQSNLIAIVRCFWFTVESGLIENHEGRKAYG 275
DB 347 QFSQDIG-----LASLGASDEEIEKLS-----VYVFTVEFGCKQNGELKAYG 390
QY 276 AVLTSSPQELGHAFIDNVRLVPLELDQIIRLPFNSTPQETLFSIRHF-----DELVELTS 331
DB 391 AGLSSYGGELLHSLSEEPVAFDPDPTAAVQYQDTYQPVYFVSESFNDKDKLRNYAS 450
QY 332 KLE 334
DB 451 RIQ 453

RESULT 14
AAR50185
ID AAR50185 standard; protein; 498 AA.
XX AAR50185;
AC AAR50185;
DT 25-MAR-2003 (revised)
DT 04-NOV-1994 (first entry)
XX Rat tyrosine hydroxylase.

XX Human; rat; tyrosine hydroxylase; TH; substitution; N-terminal;
XX enhanced enzymatic activity; wild-type; transformation; dopamine;
XX Parkinsons disease; Alzheimers disease; brain; encapsulation;
XX selectively permeable; polymer capsule; antibody.
XX Rattus rattus.
OS US5300436-A.
FN 05-APR-1994.
PD 26-JAN-1993; 93US-00009075.
PR 13-MAR-1991; 91US-00669446.
XX (UNYV) UNIV NEW YORK STATE.
XX Friedhoff AJ, Filer D, Goldstein M, Wu J;

XX⁶ WPI; 1994-125849/15.

XX Modified DNA encoding variant tyrosine hydroxylase with an N-terminal

XX amino acid substn. and cells contg. it - can be used to treat diseases

PT associated with defective function of the enzyme, e.g. Parkinson's

PT disease or Alzheimer's disease.

XX

XX Disclosure; Fig 5; 27pp; English.

XX

XX This sequence represents rat tyrosine hydroxylase (TH) and may be used in

CC the production of the variant TH molecules of the invention. These

CC variants contain amino acid substitutions in the N-terminal portion of

CC the molecule, and in the rat molecule these substitutions are pref. at

CC positions Ser8, Ser19, Ser31, Arg38, Ser40, Glu43 or Arg46. Variants

CC containing one or more of these amino acid substitutions, have

CC substantially enhanced enzymatic activity compared to the wild-type

CC enzyme. The most pref. substitution being Ser40 for Tyr or Leu. DNA

CC encoding the TH variants, and cells transfected with this DNA may be used

CC for treating diseases associated with defective function of TH, or

CC dopamine, eg. Parkinson's disease and Alzheimers disease, and effective

CC disorders. The cells can then be implanted into the brain or encapsulated

CC in a selectively permeable polymer capsule which allows release of the

CC cells products but protects them from attack by the hosts antibodies or

CC cells. (Updated on 25-MAR-2003 to correct PF field.)

XX

SQ Sequence 498 AA;

Query Match 11.8%; Score 222; DB 2; Length 498;

Best Local Similarity 26.7%; Pred. No. 2.8e-13;

Matches 65; Conservative 39; Mismatches 103; Indels 36; Gaps 5;

QY 112 SSFSLWKS-----YCPFFLDYLEAFGLSDFLDH-----QAVIKFELETHF 155

DB 227 AEEIATWKEVVTUKGLYATHACREHLEGFOLLERYCGYREDSPQLEDVSRFLKERTGF 286

QY 156 SYYPVSGFVAPHQVLSLLQDRYFFIASVMRTLDKDNFSLTPDLIHDLLGHVPMLLHPSFS 215

DB 287 QLRPVAGLLSAXDFLASLAFRVFOCTQYIRHASSPMHSPDPCCHELLGHVPMADRTFA 346

QY 216 EFFINMGRLEFTKVIKQVQALPSKKQRIQTLOSNIAlAIVRCFWFTVESGLIENHGRKAYG 275

DB 347 QFSQDIG-----LASLGASDEEIEKLS-----VWFVTFEGLCKQNGELKAYG 390

QY 276 AVLSSQELGHAFIDNVRLPLELDQIIRLPNTSTPQETLFSIRHF-----DELVELTS 331

DB 391 AGLSSYGEELLSLSEEPVEVRAFPDPTAAVQYQDTQVPYFVSEFNDKDKLRNYAS 450

QY 332 KLE 334

DB 451 RIQ 453

RESULT 15

ID AAR36741 standard; protein; 498 AA.

XX

XX AAR36741;

XX 20-SEP-1993 (first entry)

XX

XX Rat tyrosine hydroxylase.

XX

XX Human; rat; tyrosine hydroxylase; substitution; N-terminal; dopamine;

XX neurological disorder; antibody; variant.

XX Rattus rattus.

XX

XX Key Location/Qualifiers

XX Modified-site 8

XX Modified-site 19 /note= "Phosphorylation site"

XX Modified-site 19 /note= "Phosphorylation site"

FT Modified-site 40 /note= "Phosphorylation site"

FT Modified-site 153 /note= "Phosphorylation site"

XX

XX US5212082-A.

XX

XX 18-MAY-1993.

XX

XX 13-MAR-1991; 91US-00669446.

XX

XX 13-MAR-1991; 91US-00669446.

XX

XX (UUNY) UNIV NEW YORK STATE.

XX

XX Goldstein M, Wu J, Filer D, Friedhoff AJ;

XX

XX WPI; 1993-175456/21.

XX

XX Genetically modified tyrosine hydroxylase having increased activity -

XX used for treating neurological disorders e.g. Parkinson's and Alzheimer's

XX diseases and affective disorders.

XX

XX Disclosure; Fig 5; 20pp; English.

XX

XX The sequences given in AAR36740-41 represent the human and rat tyrosine

XX hydroxylase protien respectively. These sequences may be used in the

XX construction of a variant tyrosine hydroxylase which contains at least

XX one amino acid substitution in the N-terminal 55 amino acids. The

XX substitution corresponds to an amino acid selected from Ser8, Ser31,

XX Arg37, Arg38, Glu39, Ser40, Leu42, Ile42, Glu43, Asp44, Ala45, Arg46 or

XX Lys47. The enzymatic activity of the variant proteins is at least 3-fold

XX greater than that for the wild type protein. Cells transfected with the

XX DNA encoding these proteins may be used for treating neurological

XX disorders associated with a deficiency in tyrosine hydroxylase or

XX dopamine. These proteins may be used to generate antibodies specific for

XX the variant tyrosine hydroxylases to monitor the enzyme during a

XX treatment regimen

XX

SQ Sequence 498 AA;

Query Match 11.7%; Score 221; DB 2; Length 498;

Best Local Similarity 27.4%; Pred. No. 3.6e-13;

Matches 65; Conservative 37; Mismatches 99; Indels 36; Gaps 5;

QY 118 WKS-----YCPFFLDYLEAFGLSDFLDH-----QAVIKFELETHFSYYPVS 161

DB 233 WKEVVTUKGLYATHACREHLEGFOLLERYCGYREDSPQLEDVSRFLKERTGQLRVA 292

QY 162 GFVAPHQVLSLLQDRYFFIASVMRTLDKDNFSLTPDLIHDLLGHVPMLLHPSFSEFFINM 221

DB 293 GLLSAKDFLASLAFRVFOCTQYIRHASSPMHSPDPCCHELLGHVPMADRTFAQFSQDI 352

QY 222 GRLFTKVIKQVQALPSKKQRIQTLOSNIAlAIVRCFWFTVESGLIENHGRKAYGAVLISS 281

DB 353 G-----LASLGASDEEIEKLS-----VWFVTFEGLCKQNGELKAYGAGLSS 396

QY 282 POELGHAFIDNVRLPLELDQIIRLPNTSTPQETLFSIRHF-----DELVELTSKLE 334

DB 397 YGELLHSLSEEPVEVRAFPDPTAAVQYQDTQVPYFVSEFNDKDKLRNYASRIQ 453

Search completed: March 25, 2004, 14:12:27

Job time : 61 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 25, 2004, 14:10:43 ; Search time 22 Seconds
(without alignments)
849.482 Million cell updates/sec

Title: US-09-438-185A-1047
Perfect score: 1899
Sequence: 1 VHYCERTDPKVIKIALKL.....ESIFLYNQEKYLSGFEVLQ 362

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents_AA*
1: /cgm2_6/ptodata/2/iaa/5A_COMB.pep:*
2: /cgm2_6/ptodata/2/iaa/5B_COMB.pep:*
3: /cgm2_6/ptodata/2/iaa/6A_COMB.pep:*
4: /cgm2_6/ptodata/2/iaa/6B_COMB.pep:*
5: /cgm2_6/ptodata/2/iaa/PTCTUS_COMB.pep:*
6: /cgm2_6/ptodata/2/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|--------|-------------|--------|-------|----------------------|
| 1 | 1250.5 | 66.2 | 259 | 4 | US-09-198-452A-1121 |
| 2 | 237.5 | 12.6 | 308 | 4 | US-09-252-991A-29384 |
| 3 | 223 | 11.8 | 498 | 1 | US-08-009-075-5 |
| 4 | 220 | 11.6 | 497 | 1 | US-08-009-075-4 |
| 5 | 98.5 | 5.2 | 1024 | 3 | US-09-091-117-5 |
| 6 | 97 | 5.1 | 439 | 4 | US-09-198-452A-786 |
| 7 | 95 | 5.0 | 1226 | 2 | US-08-540-804-12 |
| 8 | 95 | 5.0 | 1226 | 2 | US-08-218-265-12 |
| 9 | 95 | 5.0 | 1226 | 3 | US-08-521-872-12 |
| 10 | 95 | 5.0 | 1226 | 3 | US-08-590-399-12 |
| 11 | 94 | 5.0 | 1512 | 4 | US-09-328-352-5163 |
| 12 | 92 | 4.9 | 396 | 4 | US-09-079-030-219 |
| 13 | 92 | 4.9 | 1056 | 4 | US-09-079-030-217 |
| 14 | 92 | 4.9 | 4536 | 4 | US-09-180-422B-27 |
| 15 | 92 | 4.9 | 4536 | 4 | US-09-079-030-1 |
| 16 | 92 | 4.9 | 4563 | 4 | US-09-108-006C-1 |
| 17 | 91.5 | 4.8 | 872 | 1 | US-08-766-014-2 |
| 18 | 91 | 4.8 | 535 | 3 | US-08-369-822C-24 |
| 19 | 91 | 4.8 | 535 | 3 | US-08-582-776C-39 |
| 20 | 91 | 4.8 | 535 | 3 | US-08-434-831B-36 |
| 21 | 91 | 4.8 | 2183 | 1 | US-08-348-891A-7 |
| 22 | 91 | 4.8 | 2183 | 1 | US-08-905-817-7 |
| 23 | 90.5 | 4.8 | 615 | 1 | US-08-484-105-6 |
| 24 | 90.5 | 4.8 | 615 | 1 | US-08-484-106-6 |
| 25 | 90 | 4.8 | 625 | 4 | US-09-134-001C-4504 |
| 26 | 90 | 4.8 | 654 | 4 | US-09-134-001C-3261 |
| 27 | 89 | 4.7 | 392 | 3 | US-08-911-853-35 |

```

28      89      4.7      392      3      US-09-479-409-35      Sequence 35, Appl
29      89      4.7      392      4      US-09-479-453-35      Sequence 35, Appl
30      87.5      4.6      308      4      US-09-134-000C-6494      Sequence 6494, Ap
31      87.5      4.6      309      3      US-08-996-338-22      Sequence 22, Appl
32      87.5      4.6      309      4      US-09-556-972-22      Sequence 22, Appl
33      87.5      4.6      521      3      US-08-996-338-20      Sequence 20, Appl
34      87.5      4.6      521      3      US-09-556-972-20      Sequence 20, Appl
35      87.5      4.6      541      1      US-08-604-333-2      Sequence 2, Appl
36      87.5      4.6      541      3      US-09-110-618-2      Sequence 2, Appl
37      87.5      4.6      541      4      US-09-173-151A-28      Sequence 28, Appl
38      87.5      4.6      541      4      US-09-578-178-2      Sequence 2, Appl
39      87.5      4.6      541      4      US-09-577-806-2      Sequence 2, Appl
40      87.5      4.6      541      4      US-09-621-503-4      Sequence 4, Appl
41      87      4.6      320      4      US-09-134-000C-4950      Sequence 4950, Ap
42      87      4.6      345      4      US-09-540-236-2938      Sequence 2938, Ap
43      87      4.6      437      4      US-09-540-236-2207      Sequence 2207, Ap
44      1651      4.6      1651      4      US-09-543-681A-6604      Sequence 6604, Ap
45      86.5      4.6      990      2      US-08-392-625-20      Sequence 20, Appl

```

ALIGNMENTS

```

RESULT 1
US-09-198-452A-1121
; Sequence 1121, Application US/09198452A
; Patent No. 6559294
; GENERAL INFORMATION:
; APPLICANT: Griflais, R.
; TITLE OF INVENTION: Chlamydia pneumoniae genomic sequence and polypeptides, fragment
; TITLE OF INVENTION: thereof and uses thereof, in particular for the diagnosis, prev
; TITLE OF INVENTION: and treatment of infection
; FILE REFERENCE: 9710-003-999
; CURRENT APPLICATION NUMBER: US/09/198,452A
; CURRENT FILING DATE: 1998-11-24
; NUMBER OF SEQ ID NOS: 6849
; SEQ ID NO 1121
; TYPE: PRT
; ORGANISM: Chlamydia pneumoniae
US-09-198-452A-1121

```

Query Match 66.2%; Score 1250.5; DB 4; Length 259;

Best Local Similarity 91.8%; Pred. No. 1e-126; 8; Indels 11; Gaps 1;

Matches 245; Conservative 3; Mismatches 8; Indels 11; Gaps 1;

```

Qy      96      VSTPFENRLWYRLSSRFSLWKSYCPREFLDYLEAFGLLSDFLDHQAVIKFFFELETHF 155
Db      4      MVSTPELTVFSMEKLLS-----KIFLDYLEAFGLLSDFLDHQAVIKFFFELETHF 52
Qy      156      SYTPVSGFVAPHOYLSLIQDRYFPIASVVRTDKNFSLTPLDHLIGHVPLWLLHPSFS 215
Db      53      SYTPVSGFVAPHOYLSLIQDRYFPIASVVRTDKNFSLTPLDHLIGHVPLWLLHPSFS 112
Qy      216      EPTFNNGRGLFTKVIKQVALPSKKQRIQTLQSNLIAIVRCFWFTVESGLIENHEGRKAYG 275
Db      113      EPTFNNGRGLFTKVIKQVALPSKKQRIQTLQSNLIAIVRCFWFTVESGLIENHEGRKAYG 172
Qy      276      AVLISSPOELGHAFIDNVRVLPLELDQIIRLPENTSTPQETLSIRHFDELVELTSKLEW 335
Db      173      AVLISSPOELGHAFIDNVRVLPLELDQIIRLPENTSTPQETLSIRHFDELVELTSKLEW 232
Qy      336      MLDQGLLESIPLYNQEKYLSGFEVLQ 362
Db      233      MLDQGLLESIPLYNQEKYLSGFEVLQ 259

```

```

RESULT 2
US-09-252-991A-29384
; Sequence 29384, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.

```

;; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
;; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
;; FILE REFERENCE: 107196.136
;; CURRENT APPLICATION NUMBER: US/09/252.991A
;; CURRENT FILING DATE: 1999-02-18
;; PRIOR APPLICATION NUMBER: US 60/074,788
;; PRIOR FILING DATE: 1998-02-18
;; PRIOR APPLICATION NUMBER: US 60/094,190
;; PRIOR FILING DATE: 1998-07-27
;; NUMBER OF SEQ ID NOS: 33142
;; SEQ ID NO 29384
;; LENGTH: 308
;; TYPE: PRT
;; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-29384

Query Match 12.6%; Score 237.5; DB 4; Length 308;
Best Local Similarity 24.6%; Pred. No. 3.2e-17;
Matches 60; Conservative 52; Mismatches 109; Indels 23; Gaps 5;
Qy 93 DMAVSTPFFNNLWYLLSSRSLWKSVCPRFLDYLEAFGLSLDFLDH-QAVIKFPEL 151
Db 59 DNGFIHPETEHOVWNTLTQKLVIEGRACQETDGLGIEOLGPHDRIPOLDSEINRVLOA 118
Qy 152 ETHSYYPVSGFVAPHQYLSILODRYFPFIASVMRTLDKNFSLTPDLIHDLLGHVPWLLH 211
Db 119 TTGWRVARVPALIPQTFELLASQQFPVATFIRTPEDLDYLOBPDIHFIEFGHCPLLTN 178
Qy 212 PSPEFFINMRLFTKVIKQVQALPSKQRIQLOSNIAlVRCFWFTVESGLIENHGR 271
Db 179 PWFAEFTHTYKGLKA-----SKEER-----VFLARLYWNTIEFGVETDQK 222
Qy 272 KAYGAVLISSPOELGHAFID---NVRVLPLELDQIIRLPNTSTPQETLFSIRHFDLVE 328
Db 223 RIYGGILSSPKETVYSLSDEPLHQAENPLE---AMETPYRIDLQLYFVLPDLKELFQ 279
Qy 329 LTSK 332
Db 280 LAQE 283

RESULT 3
US-09-075-5
; Sequence 5, Application US/08009075
; Patent No. 5300436
; GENERAL INFORMATION:
; APPLICANT: GOLDSTEIN, Menek
; APPLICANT: WU, Jing
; APPLICANT: FILER, David
; APPLICANT: FRIEDHOFF, Arnold J.
; TITLE OF INVENTION: GENETICALLY MODIFIED TYROSINE
; TITLE OF INVENTION: HYDROXYLASE AND USES THEREOF
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BROWDY and NEIMARK
; STREET: 419 Seventh Street, N.W., Suite 300
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION NUMBER: US/08/009,075
; FILING DATE: 19930126
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: TOWNSEND, GUY K.
; REGISTRATION NUMBER: 34,033
; REFERENCE/DOCKET NUMBER: GOLDSTEIN=1A

TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-628-5197
; TELEFAX: 202-737-3528
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 498 amino acids
; TYPE: AMINO ACID
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-009-075-5
Query Match 11.8%; Score 223; DB 1; Length 498;
Best Local Similarity 26.7%; Pred. No. 2.4e-15;
Matches 65; Conservative 39; Mismatches 103; Indels 36; Gaps 5;
Qy 112 SSRSLWKS-----YCPFFLDYLEARGLLSDFLDH-----QAVIKFPELTHF 155
Db 227 ABEIATWKEVYVTLKGLYATHACREHLEGFQLLERYCYGREDSPQLEDVSRFLKRTGF 286
Qy 156 SYVPVSGVAPHQYLSILODRYFPFIASVMRTLDKNFSLTPDLIHDLLGHVPWLLHPSFS 215
Db 287 QLRPVAGLLSARDFLASLAFRVFOCTQYIRHASSPMHSPEDCCHELLGHVPLADRTFA 346
Qy 216 EFFINMGLFTKVIKQVQALPSKQRIQLOSNIAlVRCFWFTVESGLIENHGRKAYG 275
Db 347 QPSODIG-----LASLGASDEEIEKLT-----VYWFTEVFGCKQNGELKAYG 390
Qy 276 AVLISSPOELGHAFIDNVRVLPLELDQIIRLPNTSTPQETLFSIRHF-----DELVELTS 331
Db 391 AGLLSYQELLHSLSEFEVRAFPDPAVQYQDYQYVYFVSEFSENDAKDLKNVAS 450
Qy 332 KLE 334
Db 451 RIQ 453

RESULT 4
US-08-009-075-4
; Sequence 4, Application US/08009075
; Patent No. 5300436
; GENERAL INFORMATION:
; APPLICANT: GOLDSTEIN, Menek
; APPLICANT: WU, Jing
; APPLICANT: FILER, David
; APPLICANT: FRIEDHOFF, Arnold J.
; TITLE OF INVENTION: GENETICALLY MODIFIED TYROSINE
; TITLE OF INVENTION: HYDROXYLASE AND USES THEREOF
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BROWDY and NEIMARK
; STREET: 419 Seventh Street, N.W., Suite 300
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION NUMBER: US/08/009,075
; FILING DATE: 19930126
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: TOWNSEND, GUY K.
; REGISTRATION NUMBER: 34,033
; REFERENCE/DOCKET NUMBER: GOLDSTEIN=1A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-628-5197
; TELEFAX: 202-737-3528
; INFORMATION FOR SEQ ID NO: 4:

```
;
; SEQUENCE CHARACTERISTICS:
;   LENGTH: 497 amino acids
;   TYPE: AMINO ACID
;   TOPOLOGY: linear
;   MOLECULE TYPE: protein
US-08-009-075-4
;
; Query Match 11.6%; Score 220; DB 1; Length 497;
; Best Local Similarity 26.7%; Pred. No. 5.1e-15;
; Matches 65; Conservative 38; Mismatches 104; Indels 36; Gaps 5;
;
QY 112 SSRESLWKS-----YCPRFDDYLEAFGLSDFLDH-----QAVIKFELETHF 155
Db 226 ABEATWKEVVTTLKGLYATHACGEHLEAFALLERFGYREDNTPQLEDVSRFLKERTGF 285
QY 156 SYYPVSGFVAPHQVLSLQDRYFPIASVMRTLDKDNFSLTPDLIHLGHVPMLLHPSPS 215
Db 286 QLRPVAGLLSARDFLASLAFRVFQCTQYIRHASSPMHSPEDFCHELLGHVPMADRTFA 345
QY 216 EFTINMGRLEFTKVIKQVALPSKKORIOTLOSNIATVRCFWFTVESGLIENHEGRKAYG 275
Db 346 QFSQDIG-----LASIGASDEEIEKLSLTS-----WFTVEFLCKQNGEVKAYG 389
QY 276 AVLSSPOELGHAFIDNVRLPLSLDQIIRLPNTSTPQETLSIRHF-----DELVELTS 331
Db 390 AGLSSYCELLHCLSEPEIRAFDPEAAAVQYQDTYQSYVFSSESFDKOKLRSYAS 449
QY 332 KLE 334
Db 450 RIQ 452
;
; RESULT 5
; US-09-091-117-5
; Sequence 5, Application US/09091117
; Patent No. 6171589
; GENERAL INFORMATION:
; APPLICANT: The University of Melbourne
; TITLE OF INVENTION: Mycoplasma Recombinant Polypeptides and
; TITLE OF INVENTION: Vaccines
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: GREENLEE, WINNER and SULLIVAN P.C.
; STREET: 5370 Manhattan Circle, Suite 201
; CITY: Boulder
; STATE: Colorado
; COUNTRY: United States of America
; ZIP: 80303
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/091,117
; FILING DATE: 12 JUNE 1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/AU96/00803
; FILING DATE: 13-DEC-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: AU FN7127
; FILING DATE: 13-DEC-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: WINNER, Ellen P.
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: +1 303 499 8080
; TELEFAX: +1 303 499 8089
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1024 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
```

```
;
; ORIGINAL SOURCE:
; ORGANISM: Mycoplasma genitalium
US-09-091-117-5
;
; Query Match 5.2%; Score 98.5; DB 3; Length 1024;
; Best Local Similarity 19.9%; Pred. No. 0.2;
; Matches 68; Conservative 41; Mismatches 134; Indels 99; Gaps 11;
;
QY 29 QNSQSLQRAYSTPYRYIILQKENKEK--QALARHKCISILEFFKNLLFVHLLSLSKN 85
Db 564 QOTDSLKNLFSV-----IGDILSETINVNKITHAVKNNELLSIVETASTLKIKHL----- 613
QY 86 QREGCSTDMAVSTPFFNRNLWYLLSSRFLSKWSYCPRRFLDYLEAFGLSDFLDHOAV 145
Db 614 -----NVQYKVLVDKFKLNSFIK-----ELLNFFPDTKDITPT 647
QY 146 IK--FFELETH-----FSYYPVSGFVAPHQVLSLQDRYFPIASVMRTLDKDNF 192
Db 648 IKKVLFESENYKTLKKYENEGFGYHWAKFIVPTGTFNSAENTFYSAI-----DKT 698
QY 193 SLTPDLIHLGHVPMLLHPSPSEFFINMGRLEF-----KVIEKQVALPSKKORIQT 244
Db 699 KSIRODLFADML--FGKSLESVNDSDSFIKINGSFTLKHYHGDNLNLLPNVHSLITKNVGYQI 757
QY 245 LOSNLIAIVRCFWFTVESGLIENHEGRKAYGAVLSSPQELG----- 286
Db 758 VNVNFHIDARLLTABLQNTVFSNPK-----PVIKSPVELSKSLFEVWKTFIFENSVNOI 810
QY 287 ----HAFIDNVRLPLELDQIIRLPNTSTPQETLSIRHED 324
Db 811 LKKEYTFKDLKFFPKADGSSRLSFLDKPQDVIPFAFVD 852
;
; RESULT 6
; US-09-198-452A-786
; Sequence 786, Application US/09198452A
; Patent No. 6559294
; GENERAL INFORMATION:
; APPLICANT: Griffiths, R.
; TITLE OF INVENTION: Chlamydia pneumoniae genomic sequence and polypeptides, fragment
; TITLE OF INVENTION: thereof and uses thereof, in particular for the diagnosis, prev
; TITLE OF INVENTION: and treatment of infection
; FILE REFERENCE: 9710-003-999
; CURRENT APPLICATION NUMBER: US/09/198,452A
; CURRENT FILING DATE: 1998-11-24
; NUMBER OF SEQ ID NOS: 6849
; SEQ ID NO 786
; LENGTH: 439
; TYPE: PRT
; ORGANISM: Chlamydia pneumoniae
US-09-198-452A-786
;
; Query Match 5.1%; Score 97; DB 4; Length 439;
; Best Local Similarity 22.1%; Pred. No. 0.08;
; Matches 77; Conservative 51; Mismatches 117; Indels 104; Gaps 19;
;
QY 46 RIILQKENKEKQALARHKCISIL--EFFKNLLF---VHLLSLSKNQREGCSTDMA----- 95
Db 138 RLHKNPALTHSOLVLEHTNYLKQDLKNVLFQEQFHLLAVRYNTSKHTSSLVDKLLA 197
QY 96 ----VVTSTPFNR-----NLWY--RLSSRSLWKSYPFRFFLDYLEA-----PG 134
Db 198 SYTQPISSYFSSRVERLEQISLWHQOIVNSLLEIPK---QVFLDQLTAHISGFKKQFPS 253
QY 135 LLSDFLDHQAVIKFELETH-----FSYYPVS--GVAPHQVLSLQDRYFPIAS----- 182
Db 254 ILDDL--HHEVDLLTYTSETHSSLSFSFKIAETNFHRLA-----RYKPCAATVLENMS 306
QY 183 -VMRTLDKDNF-----SLTPDLIHL-LGHVPMLLHPSPSEFFINMGRLEFTKVIKQVAL 235
Db 307 WVERTLEFCNLDRIFNTLLVLDLQEYLKQNTYTPW--LSPDESVF-----ALEKLLSS 355
QY 236 PSKKORIQTLOSNIATVRCFWFTVESGLIENHEGRKAYGAVLSSPQELGH---AFIDN 292
```

Db 356 SEQPVVQAL-----REQVLVIDFQDTKQWISFNS 390
Qy 293 VRVPLELDQIIRLPNTSPQSTLFSIRHFDLVELTSLKLEWMLDOGL 341
Db 391 LFISPRKTGSL-----FLIGDPKOSIYEWASADLPVLTAKSFSEDKQL 435

RESULT 7
US-08-540-804-12
; Sequence 12, Application US/08540804
; Patent No. 5913666
; GENERAL INFORMATION:
; APPLICANT: Young, Richard A.
; APPLICANT: Koleske, Anthony J.
; APPLICANT: Thompson, Craig M.
; APPLICANT: Chao, David M.
; TITLE OF INVENTION: No. 5919666el Factors Which Modify Gene
; TITLE OF INVENTION: Transcription and Methods of Use Thereof
; NUMBER OF SEQUENCES: 39
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
; STREET: Two Militia Drive
; CITY: Lexington
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02173
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/540,804
; FILING DATE: 11-OCT-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/521,872
; FILING DATE: 21-AUG-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/218,265
; FILING DATE: 25-MAR-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Granahan, Patricia
; REGISTRATION NUMBER: 32,227
; REFERENCE/DOCKET NUMBER: WHI94-03A2
; TELEPHONE: 617-861-6240
; TELEFAX: 617-861-9540
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1226 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-540-804-12

Query Match 5.0%; Score 95; DB 2; Length 1226;
Best Local Similarity 18.8%; Pred. No. 0.63;
Matches 98; Conservative 65; Mismatches 155; Indels 204; Gaps 25;

Qy 12 YILKIALKRLQSLRFFQNSQSLQRAYSTPYSYVRII-----LQENKKEQALAHKCI 65
Db 13 YILE---KLIFMTNHYNDSQL-RTKRQISYFKLGNCSURLINKE---IFHWLV 65
Qy 66 SILEPFKNLLF-----VHLLSKXQRCGCTDMAVST-----PFF----- 102
Db 66 EFINKVENFEPLSLHILMFWDICQIDTNPAAITSSQKPEPFLVTKITDMLHK 125
Qy 103 -----NNLWYLLSRFSL-----WKS 120
Db 126 YIVSSKSMINDENYIINDIKKNKIKILNLKILSSILKILFOQSLVFIPTSNWEI 185

Qy 121 YCRFPF-----LDYLEAFGLLSDFLDHQAIVIKFELETHESYVVSQGVAP 166
Db 186 YKPLLEFIVNADTNQNSDMKKLELISYRNESLKNSSIR-----NVIMSASNA 236
Qy 167 HQYLSLQDRYFFPIASVM-----RTLDKDN-----FSLTPDLIHLGLHVP 207
Db 237 DFQJITVYCKQPKLSQIQLNCIDTQTKLLD-DNPTFDMVTVYDQNPMTMKIQLIL 295
Qy 208 WLLHPS--PSEFFINMGRFLTKEVKQALPSKQRIQTLOSNI--IAIVRCENFTVESG 263
Db 296 WSIHPSRQFDHYESN-----QLVAKLLLL-----RINSTDEDLHEFQIDDAWSLVFQ- 343
Qy 264 LIENHEGRK-----AYGAV-----LISSP-----QELGHAFID-- 291
Db 344 LAKNFSQAKRWVSYMPSLYRLNLITYGIKVPYVIRKLISGLLYLQDSNDKFFHVQ 403
Qy 292 ---NVRVPLELDQIIRLPNTSPQSTLFSIRHFDLVELTSLK----- 333
Db 404 LLINKISPLMKSQYNNMVRNWEYDVKFEIFNFDQVETBQIKWRILSDITNLQLS 463
Qy 334 -----EWMLD---OGLLESIPLYNQEKYLSGFEVLC 361
Db 464 KTPLSIKIMVAEWYLSHLCGLSSV---NRTVLLKIFKIFC 502

RESULT 8
US-08-218-265-12
; Sequence 12, Application US/08218265
; Patent No. 5922585
; GENERAL INFORMATION:
; APPLICANT: Young, Richard A.
; APPLICANT: Koleske, Anthony J.
; APPLICANT: Thompson, Craig M.
; TITLE OF INVENTION: No. 5922585el Factors Which Modify Gene
; TITLE OF INVENTION: Transcription and Methods of Use Thereof
; NUMBER OF SEQUENCES: 35
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
; STREET: Two Militia Drive
; CITY: Lexington
; STATE: MA
; COUNTRY: US
; ZIP: 02173
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/218,265
; FILING DATE: 25-MAR-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Granahan, Patricia
; REGISTRATION NUMBER: 32,227
; REFERENCE/DOCKET NUMBER: WHI94-03
; TELEPHONE: 617-861-6240
; TELEFAX: 617-861-9540
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1226 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-218-265-12

Query Match 5.0%; Score 95; DB 2; Length 1226;
Best Local Similarity 18.8%; Pred. No. 0.63;
Matches 98; Conservative 65; Mismatches 155; Indels 204; Gaps 25;

Qy 12 YILKIALKRLQSLRFFQNSQSLQRAYSTPYSYVRII-----LQENKKEQALAHKCI 65

Db 13 YILB---KLIPDMTHYNSQOL--RTWKQISYFLKLLGNCYSRLINKE---IFHHWLV 65
QY 66 SILEFFKNLFF---VHLLSKNQREGCSTDMAVST-----PFF----- 102
Db 66 EFINKMENFELPLSLHILMIFWNDICQIDTNPVAAITSSQKEPFFLVTKITDMLLHK 125
QY 103 -----NENLWYLLSSRFSL-----WKS 120
Db 126 YIVVSSKSMINDENYIINDIKONNKIKLNLKILSSLLIKIFQBSLEVFIFPTSWEI 185
QY 121 YCPREF-----LDYLEAFGLSDFLDHOAVIKFPELETHFSYYPVSGFVAP 166
Db 186 YKPLLEFIVSNADTNQNSDMKKLELISYRNESLKNSSIR-----NVIMSASNA 236
QY 167 HOYLSLQDRYPIASVM-----RTLDKN-----FSLTPDLIHDLGHVP 207
Db 237 DFQUTIVTCKQPKLSCIQNLNCIDTQFTKLLD--DNPTFDMPTVYDQNPMTMKIQLIL 295
QY 208 WLLHPS--FSEFFINMGRFLTQVKVQALPSKKQRIQTLOSNI--IAIVRCFTWTVESG 263
Db 296 WSIHPSRQFDHYESN-----QLVAKLLLL-----RINSTDEDLHEFQIEDAIVSLVFQ- 343
QY 264 LIENHEGRK-----AYGAV-----LISSP-----QELGHAFID-- 291
Db 344 LAKNFSQKRVYSYMPSPSLYELLNITYGIKVTYIRKLSSGLLYLQDSNDKRFVHVQ 403
QY 292 ---NVRVLPLELDQIIRLPFNSTPQETLFSIRHFDLVELTSKL----- 333
Db 404 LLINKISPLMKSQYNMVLNVMYDVKFEIFNFQDLVEIQIKMRILSNDITNLQLS 463
QY 334 -----EWMLD---OGLLESIPLYNQEKYLSGFEVLC 361
Db 464 KTLPSIKIMVAEWYLSHLCSGILSV---NRTVLLKIKIFC 502

RESULT 9

US-08-521-872-12
; Sequence 12, Application US/08521872
; Patent No. 6015682
; GENERAL INFORMATION:
; APPLICANT: Young, Richard A.
; APPLICANT: Koleske, Anthony J.
; APPLICANT: Thompson, Craig M.
; APPLICANT: Chao, David M.
; TITLE OF INVENTION: No. 6015682el Factors Which Modify Gene
; TITLE OF INVENTION: Transcription and Methods of Use Therefor
; NUMBER OF SEQUENCES: 37
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
; STREET: Two Militia Drive
; CITY: Lexington
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02173
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/521,872
; FILING DATE: 31-AUG-1995
; CLASSIFICATION: 436
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: US 08/218,265
; FILING DATE: 25-MAR-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Granahan, Patricia
; REGISTRATION NUMBER: 32,227
; REFERENCE/DOCKET NUMBER: WHI94-03A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-861-6240
; TELEFAX: 617-861-9540

; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1226 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-521-872-12
Query Match 5.0%; Score 95; DB 3; Length 1226;
Best Local Similarity 18.8%; Pred. No. 0.63;
Matches 98; Conservative 65; Mismatches 155; Indels 204; Gaps 25;
QY 12 YILKIALKROSLSLFFQNSQSLQRAYSTPYSYVRII-----LOKENKEKOALAHKCI 65
Db 13 YILE---KLIPDMTHYNSQOL--RTWKQISYFLKLLGNCYSRLINKE---IFHHWLV 65
QY 66 SILEFFKNLFF---VHLLSKNQREGCSTDMAVST-----PFF----- 102
Db 66 EFINKMENFELPLSLHILMIFWNDICQIDTNPVAAITSSQKEPFFLVTKITDMLLHK 125
QY 103 -----NENLWYLLSSRFSL-----WKS 120
Db 126 YIVVSSKSMINDENYIINDIKONNKIKLNLKILSSLLIKIFQBSLEVFIFPTSWEI 185
QY 121 YCPREF-----LDYLEAFGLSDFLDHOAVIKFPELETHFSYYPVSGFVAP 166
Db 186 YKPLLEFIVSNADTNQNSDMKKLELISYRNESLKNSSIR-----NVIMSASNA 236
QY 167 HOYLSLQDRYPIASVM-----RTLDKN-----FSLTPDLIHDLGHVP 207
Db 237 DFQUTIVTCKQPKLSCIQNLNCIDTQFTKLLD--DNPTFDMPTVYDQNPMTMKIQLIL 295
QY 208 WLLHPS--FSEFFINMGRFLTQVKVQALPSKKQRIQTLOSNI--IAIVRCFTWTVESG 263
Db 296 WSIHPSRQFDHYESN-----QLVAKLLLL-----RINSTDEDLHEFQIEDAIVSLVFQ- 343
QY 264 LIENHEGRK-----AYGAV-----LISSP-----QELGHAFID-- 291
Db 344 LAKNFSQKRVYSYMPSPSLYELLNITYGIKVTYIRKLSSGLLYLQDSNDKRFVHVQ 403
QY 292 ---NVRVLPLELDQIIRLPFNSTPQETLFSIRHFDLVELTSKL----- 333
Db 404 LLINKISPLMKSQYNMVLNVMYDVKFEIFNFQDLVEIQIKMRILSNDITNLQLS 463
QY 334 -----EWMLD---OGLLESIPLYNQEKYLSGFEVLC 361
Db 464 KTLPSIKIMVAEWYLSHLCSGILSV---NRTVLLKIKIFC 502

RESULT 10

US-08-590-399-12
; Sequence 12, Application US/08590399
; Patent No. 6214588
; GENERAL INFORMATION:
; APPLICANT: Young, Richard A.
; APPLICANT: Koleske, Anthony J.
; APPLICANT: Thompson, Craig M.
; APPLICANT: Chao, David M.
; TITLE OF INVENTION: No. 6214588el Factors Which Modify Gene
; TITLE OF INVENTION: Transcription and Methods of Use Therefor
; NUMBER OF SEQUENCES: 39
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
; STREET: Two Militia Drive
; CITY: Lexington
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02173
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30


```
;
;
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US 08/590,399
; FILING DATE: 26-JAN-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/540,804
; FILING DATE: 11-OCT-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/521,872
; FILING DATE: 31-AUG-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/218,265
; FILING DATE: 25-MAR-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Granahan, Patricia
; REGISTRATION NUMBER: 32,227
; REFERENCE/DOCKET NUMBER: WH194-03A3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-861-6240
; TELEFAX: 617-861-9540
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1226 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-590-399-12

Query Match          5.0%; Score 95; DB 3; Length 1226;
Best Local Similarity 18.0%; Pred. No. 0.63;
Matches 98; Conservative 65; Mismatches 155; Indels 204; Gaps 25;

Qy 12 YILKIALKQSLFQNSQS-QRAYSTPYSYRII-----LQENKEKQALARKHCI 65
Db 13 YILE---KLIPDWNHNSQQL-RWKQGISFLKLLGNCVSLRNKE---IFHHMLV 65
Qy 66 SILEFFKNLFL---VHLSLSKNQREGCGSTDMVVST-----PFF----- 102
Db 66 EFINKMENFEFLSLHILMIFWNCIDTAPVAATITSSQKEPFLVTKITDMLHK 125
Qy 103 -----NRNLVRLSSRPSL-----RNLVRLSSRPSL-----WKS 120
Db 126 YIVSSSKSMINDENYIINDIKNNKIKNLKLSLILKIPQOSLEVFIPFSNWL 185
Qy 121 YCPFRFF-----LDYLEAGLLSDFLDHOAVIKFPELTHESYYPVSGFVAP 166
Db 186 YKPLPFIIVGNADTNQNSDMKKKLELISYRNESLKNNSIR-----NVIMSASNA 236
Qy 167 HQVLSLLQDRYFIASVM-----RTLDKDN-----FSLTDLHLLGHVP 207
Db 237 DFQITVITCKQPKLSCIQNCIDTQTKLLD-DNPTFDPWYVDQNPHTWHKIQLIL 295
Qy 208 WLLHPS--FSEFFTNMGRLETKVIERQVALPSKKQRIQTLOSNI--IAIVRCFWFTVESG 263
Db 296 WSIHPSQFDHYESN-----QLVAKLLL-----RINSTDLDHFEQIEDALWSLVFO- 343
Qy 264 LIENHEGRK-----AYGAV-----LISSP-----QELGHAFID-- 291
Db 344 LAKNFSQAKRVSYMMPSLYRLNLILITYGIKVPYIRKLISGLLYLQDSNDKXFVHQ 403
Qy 292 ---NVRVLPLELDQIIRLPNTSTPQETLPSIRHDELVELTSLK----- 333
Db 404 LLINKISPLMKSQYNMVLNRNMEYDVKFYEINFDFQDLQVEITQIKMRLISNDITNLQLS 463
Qy 334 -----EWMLD---QGLLESIFLYNOEKYLSGFEVLC 361
Db 464 KPLSLIKIMVAEWYLSHLCSGILSSV---NRTVLLKIFKIFC 502

RESULT 11
; US-09-328-352-5163
; Sequence 5163, Application US/09328352
; Patent No. 6562958

; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; FILE REFERENCE: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS
; CURRENT APPLICATION NUMBER: US/09/328,352
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 5163
; LENGTH: 1512
; TYPE: PRT
; ORGANISM: Acinetobacter baumannii
; US-09-328-352-5163

Query Match          5.0%; Score 94; DB 4; Length 1512;
Best Local Similarity 17.9%; Pred. No. 1.1;
Matches 64; Conservative 78; Mismatches 137; Indels 78; Gaps 18;

Qy 24 LSLFFQNSQSQRAYSTPYSYRIIIOKENKEKQALARKHCIISILEFFKNLL-FVHLLSL 82
Db 531 LDIFLEAEAEELLEGIDTDLNIW--VGEQEN-----FAALNNLMRYLHKLKG 574
Qy 83 SKQREGCGSTDMVVSTPFFENNLWYRLSSRPSLWKSYPFRFLDYLEAFGLSDFLDH 142
Db 575 GANMVQ--ATYLGILIAHEL--ESIYERLIQQLVATSD-----LIDFIRL--VQDDLADR 623
Qy 143 QAVIKFPELETHESYYPVSGFVAPHOVLSLL---QDRYFFIASVMRTLDKDN--FSLTPD 197
Db 624 LQIMREQQLD-----YAAPYTINALKRAQNSNFQPLSVDAFDTSEIFSEQEV 673
Qy 198 LIHLLGHVWLLHPSFSEFFTNMGRLEF-TKV-----IEKVALPSKKQRIQTLOSNI 250
Db 674 ISEVIIDIEIPVELEPALAELEAHHDQVDFATVELATPVEVITSVTSQENRVAANSODIE 733
Qy 251 AIVACFWFTVESGLIENHEG-----RKAYGAVLISSPQELGHAFIDNVRVLPLELDQI 303
Db 734 AVVEQTLEAEATELELEWAEISLLKQWFEQRTNRSILLQL-QRAVHSLKGGARMVGLQVQA 792
Qy 304 IRLPNTSTPQETLPSIRH-----DELVELTSLKLEWMLD-----QGLLESI 345
Db 793 IAYQLENAFEQ---FALHFNNSNIYDHLLE--SAIAWLKDAIFNHNHYQHPDGLQQSL 844

RESULT 12
; US-09-079-030-219
; Sequence 219, Application US/09079030
; Patent No. 6635623
; GENERAL INFORMATION:
; APPLICANT: Guevera, Jr., Juan G.
; APPLICANT: Hooogeveen, Ron C.
; APPLICANT: Moore, Paul J.
; TITLE OF INVENTION: LIPOPROTEINS AS NUCLEIC ACID DELIVERY
; NUMBER OF SEQUENCES: 229
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Arnold, White & Durkee
; STREET: P.O. Box 4433
; CITY: Houston
; STATE: Texas
; COUNTRY: USA
; ZIP: 77210
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/079,030
; FILING DATE: Concurrently Herewith
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: McMillian, Nabeela R.
; REGISTRATION NUMBER: P-43,363
```

REFERENCE/DOCKET NUMBER: ARAG:003
TELEPHONE: 512/418-3000
TELEFAX: 512/474-7577
INFORMATION FOR SEQ ID NO: 219:
SEQUENCE CHARACTERISTICS:
LENGTH: 396 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
US-09-079-030-219

Query Match 4.9%; Score 92; DB 4; Length 396;
Best Local Similarity 20.7%; Pred. No. 0.24; Indels 68; Gaps 12;
Matches 61; Conservative 48; Mismatches 118; Indels 68; Gaps 12;

QY 3 YCERLDPKYLKIALKRLQSLFFQNSQSLQRAYSTPYSYR---IILQKENKEQAL 59
Db 44 YTREELCTMFIREVGTVLQVYKVNHSSEIL-----PSYQDLVITLPPF-----L 90
QY 60 ARHKCISILEFFKNLLFVHLLSLKNOREGCSTDMVSTPFFNRNLW-----YRLSS 113
Db 91 RKHLIDVISMYREL-----LKDLSKEAQEVFKAIQSLKTTEVL-RNLQDLQIFOLIED 145
QY 114 RFLWKSYPFRFFLDYL--EAFGLSDFLDQVAVIKFEFELETHFSYYPVSGFVAP----- 166
Db 146 NIKOLKEMKFTYLYNQDEINTIFNDYIPY--VFKLLKENCLNLHKNFEFIQNELQEA 203
QY 167 -----HQYLSLQDRYFPFIASVMRT-----LDKNFSLTPDLI----- 199
Db 204 SOELQIQHYIMALREVEYFDPISVGVTVKYVELEKIVSLIKNLLVALKDFHSEYIVSAS 263
QY 200 ---HDLGHVPWLLHPSPSEFFINMGRFLTQVIEKVALPSKKQRIOTLQSNLIA 251
Db 264 NFTSOLSSQVEQFLHNTIQEYLSILTDPDGKGKKEIAELSAQAIEI--IKSQAIA 316

RESULT 13
US-09-079-030-217
Sequence 217, Application US/09079030
Patent No. 6635623
GENERAL INFORMATION:
APPLICANT: Guevera, Jr., Juan G.
APPLICANT: Hoogeveen, Ron C.
APPLICANT: Moore, Paul J.
TITLE OF INVENTION: LIPOPROTEINS AS NUCLEIC ACID DELIVERY
TITLE OF INVENTION: VECTORS FOR TRANSFECTION OF EUKARYOTIC CELLS
NUMBER OF SEQUENCES: 229
CORRESPONDENCE ADDRESS:
ADDRESSEE: Arnold, White & Durkee
STREET: P.O. Box 4433
CITY: Houston
STATE: Texas
COUNTRY: USA
ZIP: 77210
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/079,030
FILING DATE: Concurrently Herewith
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: McMillian, Nabeela R.
REGISTRATION NUMBER: P-43,363
REFERENCE/DOCKET NUMBER: ARAG:003
TELEPHONE: 512/418-3000
TELEFAX: 512/474-7577
INFORMATION FOR SEQ ID NO: 217:
SEQUENCE CHARACTERISTICS:

LENGTH: 1056 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
US-09-079-030-217

Query Match 4.9%; Score 92; DB 4; Length 1056;
Best Local Similarity 20.7%; Pred. No. 1.1; Indels 68; Gaps 12;
Matches 61; Conservative 48; Mismatches 118; Indels 68; Gaps 12;

QY 3 YCERLDPKYLKIALKRLQSLFFQNSQSLQRAYSTPYSYR---IILQKENKEQAL 59
Db 704 YTREELCTMFIREVGTVLQVYKVNHSSEIL-----PSYQDLVITLPPF-----L 750
QY 60 ARHKCISILEFFKNLLFVHLLSLKNOREGCSTDMVSTPFFNRNLW-----YRLSS 113
Db 751 RKHLIDVISMYREL-----LKDLSKEAQEVFKAIQSLKTTEVL-RNLQDLQIFOLIED 805
QY 114 RFLWKSYPFRFFLDYL--EAFGLSDFLDQVAVIKFEFELETHFSYYPVSGFVAP----- 166
Db 806 NIKOLKEMKFTYLYNQDEINTIFNDYIPY--VFKLLKENCLNLHKNFEFIQNELQEA 863
QY 167 -----HQYLSLQDRYFPFIASVMRT-----LDKNFSLTPDLI----- 199
Db 864 SOELQIQHYIMALREVEYFDPISVGVTVKYVELEKIVSLIKNLLVALKDFHSEYIVSAS 923
QY 200 ---HDLGHVPWLLHPSPSEFFINMGRFLTQVIEKVALPSKKQRIOTLQSNLIA 251
Db 924 NFTSOLSSQVEQFLHNTIQEYLSILTDPDGKGKKEIAELSAQAIEI--IKSQAIA 976

RESULT 14
US-09-180-422B-27
Sequence 27, Application US/09180422B
Patent No. 6444644
GENERAL INFORMATION:
APPLICANT: BRUCKDORFER, KARL R
ETTELAE, CAMILLE
TITLE OF INVENTION: ANTICOAGULANT PEPTIDE FRAGMENTS DERIVED
FROM APOLIPOPROTEIN B-100
NUMBER OF SEQUENCES: 27
CORRESPONDENCE ADDRESS:
ADDRESSEE: NIXON & VANDERHUE, P.C.
STREET: 1100 NORTH GLEBE ROAD, 8TH FLOOR
CITY: ARLINGTON
STATE: VA
COUNTRY: USA
ZIP: 22201
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/180,422B
FILING DATE: 07-Dec-1998
CLASSIFICATION: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: SADOFF, B.J.
REGISTRATION NUMBER: 36663
REFERENCE/DOCKET NUMBER: 117-268
TELECOMMUNICATION INFORMATION:
TELEPHONE: 7038164000
TELEFAX: 7038164100
INFORMATION FOR SEQ ID NO: 27:
SEQUENCE CHARACTERISTICS:
LENGTH: 4536 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 27:
US-09-180-422B-27

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 25, 2004, 14:13:29 ; Search time 46 Seconds
(without alignments)
2059.165 Million cell updates/sec

Title: US-09-438-185a-1047
Perfect score: 1889
Sequence: 1 VHCERTLPKYLKALKU.....ESIPLYNQKYLSGREVLQ 362

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1065169 seqs, 261661801 residues

Total number of hits satisfying chosen parameters: 1065169

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:*

1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep.*
10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
17: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
18: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|--------|-------------|--------|----|--------------------|
| 1 | 1889 | 100.0 | 362 | 15 | US-10-312-273-201 |
| 2 | 1250.5 | 66.2 | 259 | 15 | Sequence 201, App |
| 3 | 231.5 | 12.3 | 444 | 14 | US-10-289-762-1121 |
| 4 | 231.5 | 12.3 | 466 | 14 | Sequence 1121, App |
| 5 | 226.5 | 12.0 | 434 | 15 | US-10-097-340-320 |
| 6 | 223 | 11.8 | 338 | 15 | Sequence 320, App |
| 7 | 223 | 11.8 | 338 | 15 | Sequence 322, App |
| 8 | 223 | 11.6 | 338 | 15 | Sequence 1, Appl |
| 9 | 220 | 11.5 | 484 | 13 | US-10-363-474-1 |
| 10 | 216.5 | 11.5 | 485 | 13 | Sequence 22, Appl |
| 11 | 216.5 | 11.5 | 486 | 13 | Sequence 26, Appl |
| 12 | 216.5 | 11.5 | 490 | 13 | US-10-408-456-22 |
| 13 | 216.5 | 11.5 | 490 | 13 | US-10-408-456-26 |
| 14 | 206.5 | 10.9 | 532 | 9 | US-10-408-456-28 |
| 15 | 206.5 | 10.9 | 532 | 9 | US-10-408-456-24 |
| | | | | | US-10-154-674-8 |
| | | | | | Sequence 8, Appl |
| | | | | | Sequence 6, Appl |
| | | | | | Sequence 4, Appl |
| | | | | | Sequence 2, Appl |
| | | | | | Sequence 10768, A |
| | | | | | Sequence 313, App |

| | | | | | | |
|----|-------|------|------|----|----------------------|-------------------|
| 16 | 206.5 | 10.9 | 532 | 10 | US-09-963-693-313 | Sequence 313, App |
| 17 | 206.5 | 10.9 | 575 | 15 | US-10-369-493-5454 | Sequence 5454, Ap |
| 18 | 196.5 | 10.4 | 297 | 14 | US-10-138-970A-2 | Sequence 2, Appl |
| 19 | 178.5 | 9.4 | 404 | 15 | US-10-369-493-5320 | Sequence 5320, Ap |
| 20 | 175.5 | 9.3 | 457 | 15 | US-10-369-493-5501 | Sequence 5501, Ap |
| 21 | 104.5 | 5.5 | 2331 | 15 | US-10-353-856-26 | Sequence 26, Appl |
| 22 | 100.5 | 5.3 | 2212 | 15 | US-10-353-856-18 | Sequence 18, Appl |
| 23 | 100.5 | 5.3 | 2212 | 15 | US-10-353-856-36 | Sequence 36, Appl |
| 24 | 100.5 | 5.3 | 2212 | 15 | US-10-353-856-45 | Sequence 45, Appl |
| 25 | 99.5 | 5.2 | 1846 | 15 | US-10-104-047-3488 | Sequence 3488, Ap |
| 26 | 98 | 5.2 | 1846 | 15 | US-10-369-493-4974 | Sequence 4974, Ap |
| 27 | 97 | 5.1 | 439 | 15 | US-10-289-762-786 | Sequence 786, App |
| 28 | 96.5 | 5.1 | 512 | 12 | US-10-282-122A-54841 | Sequence 54841, A |
| 29 | 96 | 5.1 | 486 | 15 | US-10-369-493-2383 | Sequence 2383, Ap |
| 30 | 96 | 5.1 | 486 | 15 | US-10-369-493-2388 | Sequence 2388, Ap |
| 31 | 95.5 | 5.1 | 1564 | 16 | US-10-389-566-1234 | Sequence 1234, Ap |
| 32 | 95 | 5.0 | 1226 | 15 | US-10-369-493-1533 | Sequence 1533, Ap |
| 33 | 95 | 5.0 | 1427 | 9 | US-09-801-368-354 | Sequence 354, App |
| 34 | 94.5 | 5.0 | 2212 | 15 | US-10-353-856-8 | Sequence 8, Appl |
| 35 | 94 | 5.0 | 293 | 12 | US-10-335-977-9337 | Sequence 9337, Ap |
| 36 | 93 | 4.9 | 307 | 14 | US-10-122-466A-16 | Sequence 16, Appl |
| 37 | 93 | 4.9 | 456 | 15 | US-10-104-047-2865 | Sequence 2865, Ap |
| 38 | 92.5 | 4.9 | 466 | 12 | US-10-282-122A-63428 | Sequence 63428, A |
| 39 | 92.5 | 4.9 | 653 | 12 | US-10-425-114-54035 | Sequence 54035, A |
| 40 | 92 | 4.9 | 328 | 10 | US-09-944-049-12 | Sequence 12, Appl |
| 41 | 92 | 4.9 | 427 | 9 | US-09-817-774-23 | Sequence 23, Appl |
| 42 | 92 | 4.9 | 463 | 12 | US-10-425-114-46524 | Sequence 46524, A |
| 43 | 92 | 4.9 | 523 | 15 | US-10-369-493-1732 | Sequence 1732, Ap |
| 44 | 92 | 4.9 | 591 | 12 | US-10-282-122A-46609 | Sequence 46609, A |
| 45 | 92 | 4.9 | 670 | 14 | US-10-128-714-3587 | Sequence 3587, Ap |

ALIGNMENTS

RESULT 1
US-10-312-273-201
; Sequence 201, Application US/10312273
; Publication No. US20040005667A1
; GENERAL INFORMATION:
; APPLICANT: CHIRON Spa
; TITLE OF INVENTION: IMMUNISATION AGAINST CHLAMYDIA PNEUMONIAE
; FILE REFERENCE: P025035WO
; CURRENT APPLICATION NUMBER: US/10/312, 273
; PRIOR FILING DATE: 2002-12-20
; PRIOR FILING DATE: 2000-07-03
; PRIOR APPLICATION NUMBER: 0016363.4
; PRIOR FILING DATE: 2000-07-11
; PRIOR APPLICATION NUMBER: 0017047.2
; PRIOR FILING DATE: 2000-07-11
; PRIOR APPLICATION NUMBER: 0017983.8
; PRIOR FILING DATE: 2000-07-21
; PRIOR APPLICATION NUMBER: 0019369.0
; PRIOR FILING DATE: 2000-08-07
; PRIOR APPLICATION NUMBER: 0020440.4
; PRIOR FILING DATE: 2000-08-18
; PRIOR APPLICATION NUMBER: 0022583.9
; PRIOR FILING DATE: 2000-09-14
; PRIOR APPLICATION NUMBER: 0027549.5
; PRIOR FILING DATE: 2000-11-10
; PRIOR APPLICATION NUMBER: 0031706.5
; PRIOR FILING DATE: 2000-12-22
; NUMBER OF SEQ ID NOS: 664
; SOFTWARE: SeqWin99, version 1.02
; SEQ ID NO 201
; LENGTH: 362
; TYPE: PRT
; ORGANISM: Chlamydia pneumoniae
US-10-312-273-201

Query Match 100.0%; Score 1889; DB 15; Length 362;
Best Local Similarity 100.0%; Pred. No. 2, 1e-185;
Matches 362; Conservative 0; Mismatches 0; Indels 0; Gaps 0;


```

; GENERAL INFORMATION:
; APPLICANT: John Monahan
; APPLICANT: Manjula Cannavarapu
; APPLICANT: Sebastian Hoersch
; APPLICANT: Shubhangi Kamarkar
; APPLICANT: Steve G. Kovats
; APPLICANT: Rachel E. Meyers
; APPLICANT: Michael Morrissey
; APPLICANT: Peter Olandt
; APPLICANT: Ami Sen
; APPLICANT: Peter Veiby
; APPLICANT: Gordon B. Mills
; APPLICANT: Robert C. Bast, Jr.
; APPLICANT: Karen Lu
; APPLICANT: Rosemarie Schmandt
; APPLICANT: Xumei Zhao
; APPLICANT: Karen Glatt
; TITLE OF INVENTION: Nucleic Acid Molecules and Proteins For The Identification,
; TITLE OF INVENTION: Assessment, Prevention, and Therapy of Ovarian Cancer
; FILE REFERENCE: MRI-030
; CURRENT APPLICATION NUMBER: US/10/097,340
; CURRENT FILING DATE: 2002-03-14
; PRIOR APPLICATION NUMBER: 60/276,025
; PRIOR FILING DATE: 2001-03-14
; PRIOR APPLICATION NUMBER: 60/325,149
; PRIOR FILING DATE: 2001-09-26
; PRIOR APPLICATION NUMBER: 60/276,026
; PRIOR FILING DATE: 2001-03-14
; PRIOR APPLICATION NUMBER: 60/324,967
; PRIOR FILING DATE: 2001/09/26
; PRIOR APPLICATION NUMBER: 60/311,732
; PRIOR FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: 60/325,102
; PRIOR FILING DATE: 2001-09-26
; PRIOR APPLICATION NUMBER: 60/323,580
; PRIOR FILING DATE: 2001-09-19
; NUMBER OF SEQ ID NOS: 363
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 322
; LENGTH: 466
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-097-340-322

Query Match      12.3%; Score 231.5; DB 14; Length 466;
Best Local Similarity 27.8%; Pred. No. 8.6e-15;
Matches 64; Conservative 43; Mismatches 96; Indels 27; Gaps 4;

QY 107 WYRLSSRFSLWKSVCPRFFLDYLEAFGLLSDFLDH-----QAVIKFFELETHESYY 158
DB 174 WGTVPQELNKLYPTHACR---EYLKLPFLSKYCYREDNIPQLEDVSNFLKERTGFSIR 230
QY 159 PVSGFVAPHQYLSLQDRYFFPIASVMRTLDKDNFSLTPDLIHDLGHVPMLLHPSFSEFF 218
DB 231 PVAGYLSPRDFLSGLAFRVFCHTQVVRHSSDPFVTPEDTCHELLGHVPLLAEPFSAQFS 290
QY 219 INMGRLFTKVIKQVALSKKORIQTLOSNIATVRCFWFVSGLIENHEGRKAYGAVL 278
DB 291 QEIG-----LASLGASEAVQKLT-----CYFFTFEGLCKQDQLRVFGAGL 334

Query Match      11.8%; Score 223; DB 15; Length 338;
Best Local Similarity 26.7%; Pred. No. 4.1e-14;
Matches 65; Conservative 39; Mismatches 103; Indels 36; Gaps 5;

QY 112 SSRFSLWKS-----YCPRFELDYLEAFGLLSDFLDH-----QAVIKFFELETHF 155
DB 67 ABEIATWKEVYTTKGLYATHACGHELEAFALLERFSGYREDNIPQLEDVSNFLKERTG 126
QY 156 SYFVSGFVAPHQYLSLQDRYFFPIASVMRTLDKDNFSLTPDLIHDLGHVPMLLHPSFS 215
DB 127 QLAPVAGLLSARDFLASLAFRVFCTQYIRHASSPMHSPEDCCHELLGHVPMLARPTFA 186
QY 216 EFFINMGRFLTKEVQALPSKKORIQTLOSNIATVRCFWFVSGLIENHEGRKAYG 275

; TITLE OF INVENTION: hemostasis deficiency
; FILE REFERENCE: 5013.1052
; CURRENT APPLICATION NUMBER: US/10/363,474
; CURRENT FILING DATE: 2003-07-22
; PRIOR APPLICATION NUMBER: PCT/DE01/03178
; PRIOR FILING DATE: 2001-08-27
; PRIOR APPLICATION NUMBER: DE 10043124.0
; PRIOR FILING DATE: 2000-08-31
; NUMBER OF SEQ ID NOS: 1
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 1
; LENGTH: 434
; TYPE: PRT
; ORGANISM: homo sapiens
; US-10-363-474-1

Query Match      12.0%; Score 226.5; DB 15; Length 434;
Best Local Similarity 27.4%; Pred. No. 2.6e-14;
Matches 63; Conservative 43; Mismatches 97; Indels 27; Gaps 4;

QY 107 WYRLSSRFSLWKSVCPRFFLDYLEAFGLLSDFLDH-----QAVIKFFELETHESYY 158
DB 164 WGTVPRELNKLYPTHACR---EYLRLPLLSKYCYREDNIPQLEDVSNFLKERTGFSIR 220
QY 159 PVSGFVAPHQYLSLQDRYFFPIASVMRTLDKDNFSLTPDLIHDLGHVPMLLHPSFSEFF 218
DB 221 PVAGYLSPRDFLSGLAFRVFCHTQVVRHSSDPLYTPEDTCHELLGHVPLLAEPFSAQFS 280
QY 219 INMGRLFTKVIKQVALPSKKORIQTLOSNIATVRCFWFVSGLIENHEGRKAYGAVL 278
DB 281 QEIG-----LASLGASEETVQKLT-----CYFFTFEGLCKQDQLRVFGAGL 324
QY 279 ISSPOELGHAFIDNVRLPLELDQIIRLPNTSTPQETLFSIRHFDLVE 328
DB 325 LSSISELKHLSGHAKVPFDPKICKQECCLITTFQDVYFVSESFEDAKE 374

RESULT 6
US-10-408-456-22
; Sequence 22, Application US/10408456
; Publication No. US20040013648A1
; GENERAL INFORMATION:
; APPLICANT: Oxford Biomedica (UK) Limited
; APPLICANT: KINGSMAN, et al., Alan John
; TITLE OF INVENTION: Vector System
; FILE REFERENCE: 674523-2016
; CURRENT APPLICATION NUMBER: US/10/408,456
; CURRENT FILING DATE: 2003-04-08
; PRIOR APPLICATION NUMBER: PCT/GB01/04433
; PRIOR FILING DATE: 2001-10-15
; PRIOR APPLICATION NUMBER: 0024550.6
; PRIOR FILING DATE: 2000-10-06
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 22
; LENGTH: 338
; TYPE: PRT
; ORGANISM: Homo Sapiens
; US-10-408-456-22

Query Match      11.8%; Score 223; DB 15; Length 338;
Best Local Similarity 26.7%; Pred. No. 4.1e-14;
Matches 65; Conservative 39; Mismatches 103; Indels 36; Gaps 5;

QY 112 SSRFSLWKS-----YCPRFELDYLEAFGLLSDFLDH-----QAVIKFFELETHF 155
DB 67 ABEIATWKEVYTTKGLYATHACGHELEAFALLERFSGYREDNIPQLEDVSNFLKERTG 126
QY 156 SYFVSGFVAPHQYLSLQDRYFFPIASVMRTLDKDNFSLTPDLIHDLGHVPMLLHPSFS 215
DB 127 QLAPVAGLLSARDFLASLAFRVFCTQYIRHASSPMHSPEDCCHELLGHVPMLARPTFA 186
QY 216 EFFINMGRFLTKEVQALPSKKORIQTLOSNIATVRCFWFVSGLIENHEGRKAYG 275

; TITLE OF INVENTION: Method for diagnosing neuronal diseases and for treating primary
```



```

; SEQ ID NO 6
; LENGTH: 485
; TYPE: PRT
; ORGANISM: homo sapiens
US-10-154-674-6

Query Match
  11.5%   Score 216.5;   DB 13;   Length 485;
Best Local Similarity 26.2%;   Pred. No. 3.2e-13;
Matches 67;   Conservative 42;   Mismatches 104;   Indels 43;   Gaps 4;

Qy 117 LWSYCPRFLLDYLEAFGLLSDFLDH-----QAVIKFFELETHFSYYPVSGFVAPHQ 168
Db 222 LSKLYPTHACREYLNKFPPLTKYCGYREDNVPOLEDVSMFLKRSGETVVPVAGYLSPRD 281
Qy 169 YLSLLQDRYPPIASVNRITLKDNYFSLTPDLIHLGHVPWLLHPSGFEFFINMGRLETKV 228
Db 282 FLAGLAVRPHCTQYIRHGSDPLVTPDPDTCHELLGHVPLADPKPAQFSQEIG----- 335
Qy 229 IEKVOALPSKKRIQTLQSNLIAIVRCFTWTVSGLIENHEGRKAYCAVLISSPQELGHA 288
Db 336 ---LASIGASDEDVQKLAT-----CYFTIIFGLCKQEGQLRAYGAGLLSSIGELKHA 385
Qy 289 FIDNRVRLPELDOIIRLPNTTPOBTLFSIRHFDLVE-----L 329
Db 386 LSDKACVKAFDPKTTCLQCELIITFQAYFVSSFEAKKMRDPAKSITRPFVSYPFY 445
Qy 330 TSKLEWMLDOGLLESI 345
Db 446 TQSIETLKDTRSIENV 461

RESULT 12
US-10-154-674-4
; Sequence 4, Application US/10154674
; Publication No. US20020192694A1
; GENERAL INFORMATION:
; APPLICANT: Yu, Xuanchuan
; APPLICANT: Miranda, Maricar
; APPLICANT: Hu, Yi
; TITLE OF INVENTION: No. US20020192694A1el Human Hydroxylases and Polynucleotides Enc
; FILE REFERENCE: LEX-0352-USA
; CURRENT APPLICATION NUMBER: US/10/154,674
; CURRENT FILING DATE: 2002-05-23
; PRIOR APPLICATION NUMBER: US 60/294,076
; PRIOR FILING DATE: 2001-05-29
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 486
; TYPE: PRT
; ORGANISM: homo sapiens
US-10-154-674-4

Query Match
  11.5%   Score 216.5;   DB 13;   Length 486;
Best Local Similarity 26.2%;   Pred. No. 3.2e-13;
Matches 67;   Conservative 42;   Mismatches 104;   Indels 43;   Gaps 4;

Qy 117 LWSYCPRFLLDYLEAFGLLSDFLDH-----QAVIKFFELETHFSYYPVSGFVAPHQ 168
Db 223 LSKLYPTHACREYLNKFPPLTKYCGYREDNVPOLEDVSMFLKRSGETVVPVAGYLSPRD 282
Qy 169 YLSLLQDRYPPIASVNRITLKDNYFSLTPDLIHLGHVPWLLHPSGFEFFINMGRLETKV 228
Db 283 FLAGLAVRPHCTQYIRHGSDPLVTPDPDTCHELLGHVPLADPKPAQFSQEIG----- 336
Qy 229 IEKVOALPSKKRIQTLQSNLIAIVRCFTWTVSGLIENHEGRKAYCAVLISSPQELGHA 288
Db 337 ---LASIGASDEDVQKLAT-----CYFTIIFGLCKQEGQLRAYGAGLLSSIGELKHA 386
Qy 289 FIDNRVRLPELDOIIRLPNTTPOBTLFSIRHFDLVE-----L 329
Db 387 LSDKACVKAFDPKTTCLQCELIITFQAYFVSSFEAKKMRDPAKSITRPFVSYPFY 446

```


QY 330 TSKLEWMLDQGLLESI 345
Db 447 TQSEILKDTRSIENV 462

RESULT 13
US-10-154-674-2
; Sequence 2, Application US/10154674
; Publication No. US20020192694A1
; GENERAL INFORMATION:
; APPLICANT: Yu, Xuanchuan
; APPLICANT: Miranda, Maricar
; APPLICANT: Hu, Yi
; TITLE OF INVENTION: No. US20020192694A1 Human Hydroxylases and Polynucleotides Encoded by the Same Gene
; FILE REFERENCE: LEX-0352-USA
; CURRENT APPLICATION NUMBER: US/10/154,674
; CURRENT FILING DATE: 2002-05-23
; PRIOR APPLICATION NUMBER: US 60/294,076
; PRIOR FILING DATE: 2001-05-29
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 490
; TYPE: PRT
; ORGANISM: homo sapiens
US-10-154-674-2

Query Match 11.5%; Score 216.5; DB 13; Length 490;
Best Local Similarity 26.2%; Pred. No. 3.2e-13;
Matches 67; Conservative 42; Mismatches 104; Indels 43; Gaps 4;

QY 117 LKWSYCPFFLDYLAFLGSLDLDH-----QAVIKPFELETHFSYYPVSGFVAPHQ 168
Db 227 LSKLYPHACREYLNKPNPLITKYCYREDNVPOLEDVSMFLKERSGFTVVRVAGYLSPRD 286

QY 169 YLSIQDRYFPIASVMTLDKNSLTPLDHLHGHVPMWLLHPSFSEFFINMGLFTKV 228
Db 287 FLAGLAVRFHCTQYIHGSDPLYTEPTDCHLLGHVPLADPKPAQFSQIG----- 340

QY 229 IEKVOALPSKKQRIQTLQSNLIAIVCFWFTVSGLIENHGRKAYGAVLISSPQELGHA 288
Db 341 ---LASLGASDEDVOKLAT-----CYFTTIEFLGCKQEGQLRAYCAGLLSSIGELKHA 390

QY 289 FIDNRVLPLELDQIRLTPNTSTPQETLFSIRHDELVE-----L 329
Db 391 LSKKACVAFDPTKTCQELITTFQAYFVSSEFEAKKMDRPAKSIITRPFVSIVYFNFY 450

QY 330 TSKLEWMLDQGLLESI 345
Db 451 TQSEILKDTRSIENV 466

RESULT 14
US-10-369-493-10768
; Sequence 10768, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10/52052/B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 10768
; LENGTH: 253
; TYPE: PRT

; ORGANISM: Spingomonas aromaticivorans
US-10-369-493-10768

Query Match 10.9%; Score 206.5; DB 15; Length 253;
Best Local Similarity 26.3%; Pred. No. 1.3e-12;
Matches 61; Conservative 36; Mismatches 108; Indels 27; Gaps 4;

QY 106 LWYRLSSRSFLSKSYCPFFLDYLAFLGSLDLDHQAIVIKFELE-----THFSYYP 159
Db 9 IWELIARQWELLPGACSAFLQGLERLD-----LGRGGVDFPARLSSELGALTGWSVVP 63

QY 160 VSGFVAPHQYLSLQDRYFPIASVMTLDKNSLTPLDHLHGHVPMWLLHPSFSEFFI 219
Db 64 VPMILPDHVFVFWHLANRRFPAGNFIRTFEDYIQEPDVFHDFVGHVPMULTDTFYADYMQ 123

QY 220 NMGLFTKVIKQALPSKKQRIQTLQSNLIAIVCFWFTVSGLIENHGRKAYGAVLI 279
Db 124 EYGRAGWKAMR-----YNRLKALGALYWTYVEFGLVIEDGAPKYVGAIL 168

QY 280 SSPQELGHAFI-DNRVLPLELDQIRLTPNTSTPQETLFSIRHDELVELT 330
Db 169 SCPREAVFALEGQSPNRMILNVDVWRTDYVIDDLOPTYFVIESFADLYHOT 220

RESULT 15
US-09-205-658-313
; Sequence 313, Application US/09205658
; Patent No. US20010029617A1
; GENERAL INFORMATION:
; APPLICANT: Ruvkun, Gary
; APPLICANT: Ogg, Scott
; TITLE OF INVENTION: THERAPEUTIC AND DIAGNOSTIC TOOLS FOR
; FILE OF INVENTION: IMPAIRED GLUCOSE TOLERANCE CONDITIONS
; CURRENT APPLICATION NUMBER: US/09/205,658
; CURRENT FILING DATE: 1998-12-03
; EARLIER APPLICATION NUMBER: 08/857,076
; EARLIER FILING DATE: 1997-05-15
; EARLIER APPLICATION NUMBER: 08/888,534
; EARLIER FILING DATE: 1997-07-07
; EARLIER APPLICATION NUMBER: US98/10080
; EARLIER FILING DATE: 1998-05-15
; NUMBER OF SEQ ID NOS: 328
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 313
; LENGTH: 532
; TYPE: PRT
; ORGANISM: Caenorhabditis elegans
US-09-205-658-313

Query Match 10.9%; Score 206.5; DB 9; Length 532;
Best Local Similarity 25.2%; Pred. No. 3.9e-12;
Matches 70; Conservative 47; Mismatches 116; Indels 45; Gaps 8;

QY 104 RNLM---YRLSSRSFLSKSYCPFFLDYLAFLGSLDLDH-----QAVIKPFELETHF 155
Db 251 RKTWGIIVYKLR---ELHKKHACKQFLDNFELLERHCGYSENNIPQLEDICKLUKATGF 307

QY 156 SVYPVSGFVAPHQYLSLQDRYFPIASVMTLDKNSLTPLDHLHGHVPMWLLHPSFS 215
Db 308 RVRPVAGYLSARDFLAGLAYRVFCTQVVRHADPFYTPEDTVHLMGHMALFADPDFA 367

QY 216 EFFINMGLFTKVIKQALPSKKQRIQTLQSNLIAIVCFWFTVSGLI----- 265
Db 368 QFSQIG-----LASLGASDEDLKLATL-----YFSEIFGLSSDDAADSVPK 411

QY 266 ---ENHGRKAYGAVLISSPQELGHAFIDNRVLPLELDQIRLTPNTSTPQETLFSIRH 322
Db 412 ENGSHREFKYGAGLLSSAGELQHAVESGATIRFDRVVEQECLITTFQSAFYFTRN 471

QY 323 FDELVELTSKLEWMLDQGLLESIPLYNQEKYLSGFEVL 360
Db 472 FEEAQQ---KLRFMTNNMKRPFIVRYN--PYTESVEVL 504

us-09-438-185a-1047.rapb

Mon Mar 29 12:10:57 2004

Search completed: March 25, 2004, 14:19:13
Job time : 47 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 25, 2004, 14:10:43 ; Search time 20 Seconds

(without alignments)
1741.067 Million cell updates/sec

Title: US-09-438-185A-1047

Perfect score: 1889

Sequence: 1 VHYCERTLPKYLKIALKL.....ESIPLYNQKYLSGFEVLQ 362

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR 78.*

1: pir1.*

2: pir2.*

3: pir3.*

4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|---------|--------------------|
| 1 | 1886 | 99.8 | 362 | E72002 | aromatic amino aci |
| 2 | 1886 | 99.8 | 362 | C86621 | aromatic amino aci |
| 3 | 256.5 | 13.6 | 289 | D82413 | phenylalanine-4-hy |
| 4 | 235.5 | 12.5 | 262 | F83535 | phenylalanine-4-hy |
| 5 | 231.5 | 12.3 | 444 | S10489 | tryptophan 5-monoo |
| 6 | 229.5 | 12.1 | 262 | A53452 | phenylalanine hydr |
| 7 | 229.5 | 12.1 | 444 | S81199 | tryptophan 5-monoo |
| 8 | 229.5 | 12.1 | 491 | A28582 | tyrosine 3-monooxy |
| 9 | 227.5 | 12.0 | 444 | WHRTW | tryptophan 5-monoo |
| 10 | 226.5 | 12.0 | 447 | A34582 | tryptophan 5-monoo |
| 11 | 225.5 | 11.9 | 498 | JN0068 | tyrosine 3-monooxy |
| 12 | 223.5 | 11.8 | 481 | IS1567 | tryptophan 5-monoo |
| 13 | 223.5 | 11.8 | 498 | WHRTY | tyrosine 3-monooxy |
| 14 | 221.5 | 11.7 | 491 | I45983 | tyrosine 3-monooxy |
| 15 | 220 | 11.6 | 528 | WHHUY4 | tyrosine 3-monooxy |
| 16 | 219 | 11.6 | 579 | A55369 | tyrosine 3-monooxy |
| 17 | 214 | 11.3 | 453 | WHRTF | phenylalanine 4-mo |
| 18 | 213.5 | 11.3 | 452 | 1 WHHUF | phenylalanine 4-mo |
| 19 | 212.5 | 11.2 | 453 | S15758 | phenylalanine 4-mo |
| 20 | 211.5 | 11.2 | 491 | JL0039 | tyrosine 3-monooxy |
| 21 | 207 | 11.0 | 453 | JQ0766 | phenylalanine 4-mo |
| 22 | 206.5 | 10.9 | 575 | T34509 | hypothetical prote |
| 23 | 197 | 10.4 | 453 | A44271 | tryptophan 5-monoo |
| 24 | 194 | 10.3 | 452 | JC4888 | phenylalanine 4-hy |
| 25 | 189.5 | 10.0 | 294 | C87449 | phenylalanine 4-mo |
| 26 | 178.5 | 9.4 | 404 | T25453 | tyrosine 3-monooxy |
| 27 | 175.5 | 9.3 | 457 | T23494 | phenylalanine 4-mo |
| 28 | 109.5 | 5.8 | 1253 | T45787 | disease resistance |
| 29 | 108.5 | 5.7 | 1501 | B29813 | 174K ninaC protein |

| | | | | | | |
|----|-------|-----|------|---|--------|---------------------|
| 30 | 106.5 | 5.6 | 296 | 2 | A40996 | phenylalanine 4-mo |
| 31 | 104.5 | 5.5 | 2331 | 2 | S44054 | genome polypeptide |
| 32 | 102 | 5.4 | 563 | 2 | T38766 | probable transcrip |
| 33 | 102 | 5.4 | 1050 | 2 | C81624 | exodeoxyribonuclea |
| 34 | 100.5 | 5.3 | 1157 | 2 | S38160 | NUP133 protein - y |
| 35 | 99 | 5.2 | 924 | 2 | T09220 | exocyst complex pr |
| 36 | 98.5 | 5.2 | 584 | 2 | D81265 | hypothetical prote |
| 37 | 98.5 | 5.2 | 1024 | 2 | C64208 | hypothetical prote |
| 38 | 98 | 5.2 | 1846 | 2 | T33079 | cytochrome P450 2L |
| 39 | 97 | 5.1 | 492 | 1 | S68856 | exodeoxyribonuclea |
| 40 | 97 | 5.1 | 1050 | 2 | G8582 | exodeoxyribonuclea |
| 41 | 96.5 | 5.1 | 512 | 2 | F8520 | glucose-6-P dehydro |
| 42 | 96.5 | 5.1 | 512 | 2 | C72103 | glucose-6-P dehydro |
| 43 | 96 | 5.1 | 529 | 2 | S62468 | probable membrane |
| 44 | 96 | 5.1 | 1050 | 2 | H72041 | exodeoxyribonuclea |
| 45 | 96 | 5.1 | 2059 | 2 | T41933 | large tegument pro |

ALIGNMENTS

RESULT 1

E72002
aromatic amino acid hydroxylase - Chlamydomophila pneumoniae (strain CML029)
C:Species: Chlamydomophila pneumoniae, Chlamydia pneumoniae
C:Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 05-May-2000
C:Accession: E72002
R:Kallman, S.; Mitchell, W.; Marathe, R.; Lammel, C.; Fan, J.; Olinger, L.; Grimwood, J.
Nature Genet. 21, 385-389, 1999
A:Title: Comparative Genomes of Chlamydia pneumoniae and C. trachomatis.
A:Reference number: A72000, MUID:99206606, PMID:10192388
A:Accession: E72002
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-362 <ARN>
A:Cross-references: GB:AE001685; GB:AE001363; NID:94377378; PIDN:AAD19183.1; PID:943773
A:Experimental source: strain CML029
C:Genetics:
A:Gene: Cpa1046

Query Match 99.8%; Score 1886; DB 2; Length 362;
Best Local Similarity 99.7%; Pred. No. 2.1e-144;
Matches 361; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

| | | | |
|----|-----|--|-----|
| Qy | 1 | VHYCERTLPKYLKIALKROSLSLFFQNSQSLORAYSTPYSYRRIILOKENEKQALA | 60 |
| Db | 1 | MHYCERTLPKYLKIALKROSLSLFFQNSQSLORAYSTPYSYRRIILOKENEKQALA | 60 |
| Qy | 61 | RHKCISILEFFKNLLFVHLLSLSKNOREGCGTDMAVVSTPPFNRLWYLLSSRFLSKS | 120 |
| Db | 61 | RHKCISILEFFKNLLFVHLLSLSKNOREGCGTDMAVVSTPPFNRLWYLLSSRFLSKS | 120 |
| Qy | 121 | YCFRFFLDYLEAFGLSLDFLDHQAVTKPFEETHSYYPVSGFVAPHOYLSLLQDRYFPI | 180 |
| Db | 121 | YCFRFFLDYLEAFGLSLDFLDHQAVTKPFEETHSYYPVSGFVAPHOYLSLLQDRYFPI | 180 |
| Qy | 181 | ASVWRTLDKDNFSLTDLIHLHLGHVPWLLHPSSEFFINMGRLFTKVIKQALPSKKQ | 240 |
| Db | 181 | ASVWRTLDKDNFSLTDLIHLHLGHVPWLLHPSSEFFINMGRLFTKVIKQALPSKKQ | 240 |
| Qy | 241 | RIOTLOSGLIAVRCFWFTVESGLIENHGRKAYGAVLSSPOELGHAFIDNVRVLPLEL | 300 |
| Db | 241 | RIOTLOSGLIAVRCFWFTVESGLIENHGRKAYGAVLSSPOELGHAFIDNVRVLPLEL | 300 |
| Qy | 301 | DQIIRLPFNTSTPQETLFSIRHFDDELVELTSKLEWMLDQGLLESIPLYNQKYLSGFEVL | 360 |
| Db | 301 | DQIIRLPFNTSTPQETLFSIRHFDDELVELTSKLEWMLDQGLLESIPLYNQKYLSGFEVL | 360 |
| Qy | 361 | CQ 362 | |
| Db | 361 | CQ 362 | |

RESULT 2

C86621
aromatic amino acid hydroxylase [imported] - Chlamydomonas aeruginosa (strain J138)
C:Species: Chlamydomonas aeruginosa
C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 02-Mar-2001
C:Accession: C86621
R:Shirai, M.; Hirakawa, H.; Kimoto, M.; Tabuchi, M.; Kishi, F.; Ouchi, K.; Shiba, T.; Ishii, T.
Nucleic Acids Res. 28, 2311-2314, 2000.
A:Title: Comparison of whole genome sequences of chlamydia pneumoniae J138.
A:Reference number: A86491; MUID:20330349; PMID:10871362
A:Accession: C86621
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-362 <STO>
A:CROSS-references: GB:BA000008; NID:9879419; PIDN:BAA95253.1; GSPDB:GN00142
A:Experimental source: strain J138
C:Genetics:
A:Gene: CP1046

Query Match 99.8%; Score 1886; DB 2; Length 362;
Best Local Similarity 99.7%; Pred. No. 2.1e-144;
Matches 361; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 VHCERTLDPKYILKIALKQSLFFQNSQSLQRAYSTPSYYRIILQKENKEQALA 60
DB 1 MHVCERTLDPKYILKIALKQSLFFQNSQSLQRAYSTPSYYRIILQKENKEQALA 60
QY 61 RHKICISILEFFKNLLFVHLISLKNQREGGSDMAVSTPFFNRLWYRLLSRFSWLKS 120
DB 61 RHKICISILEFFKNLLFVHLISLKNQREGGSDMAVSTPFFNRLWYRLLSRFSWLKS 120
QY 121 YCPREFLDYLEAFGLSDFLDHOAVIKFPELETHFSYPSVGVAPHOYLSLLQDRYFFI 180
DB 121 YCPREFLDYLEAFGLSDFLDHOAVIKFPELETHFSYPSVGVAPHOYLSLLQDRYFFI 180
QY 181 ASVNRITLDKDNFSLTPDLIHLGHVPLMLHPSPFNNMGLFTKVIKQVALSKQ 240
DB 181 ASVNRITLDKDNFSLTPDLIHLGHVPLMLHPSPFNNMGLFTKVIKQVALSKQ 240
QY 241 RIQTLQNLIAIVRCFWFTVESGLIENHEGRKAYGAVLISSPOELGHAFIDNVRVLPLEL 300
DB 241 RIQTLQNLIAIVRCFWFTVESGLIENHEGRKAYGAVLISSPOELGHAFIDNVRVLPLEL 300
QY 301 DOIIRLPNTSTPQETILFSIRHFDVLTSLKLEWMLDQGLLESIPLYNOEKYLSGFEVL 360
DB 301 DOIIRLPNTSTPQETILFSIRHFDVLTSLKLEWMLDQGLLESIPLYNOEKYLSGFEVL 360
QY 361 CQ 362
DB 361 CQ 362

RESULT 3

D82413
phenylalanine-4-hydroxylase VCA0828 [imported] - Vibrio cholerae (strain N16961 serogroup O1)
C:Species: Vibrio cholerae
C:Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 02-Feb-2001
C:Accession: D82413
R:Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.J.; Chardson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragoi, I.; Sellers, P.; R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.
Nature 406, 477-483, 2000
A:Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.
A:Reference number: A82035; MUID:20406833; PMID:10952301
A:Accession: D82413
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-289 <HEI>
A:CROSS-references: GB:AB004410; GB:AB003853; NID:9656244; PIDN:AAF96726.1; GSPDB:GN00142
A:Experimental source: serogroup O1; strain N16961; biotype El Tor
C:Genetics:
A:Gene: VCA0828
A:Map position: 2

Query Match 13.6%; Score 256.5; DB 2; Length 289;

Best Local Similarity 25.6%; Pred. No. 3.2e-13;
Matches 65; Conservative 56; Mismatches 108; Indels 25; Gaps 6;

QY 106 LWYLLSSRFLSKSYCPRFFLDYLEAFGLSDFLDHOAVI-KFFELETHFSYPSVGFV 164
DB 49 VVHELITRQEVVVKTRACQYLDGLNMLNLTORLPOLPEINRVLORETQGWEPVALL 108
QY 165 APHOYLSLQDRYPIASVMTLDKONFSLTPDLIHLGHVPMWLLHPSSEFFINMGR 224
DB 109 SDFRFFALLADKFPVATFLRREEDYLDQEPDFHEVYGHGCMALHDPDFAATHVYGL 168
QY 225 FTKVIEKVALPSKKQRIQTLQSNLIAIVRCFWFTVESGLIENHEGRKAYGAVLISSPQE 284
DB 169 GAKATPKERSY-----LRLYNTVEFGLVQEQGQTKYGGGILSPGE 212
QY 285 LGHAFIDNV-RVLPLELDQIIRLPNTSTPQETILFSIRHFDVLTSLK-----LEWM-LD 338
DB 213 TLYASESTIPKREPFDIMQVLRTPYRIDIMQYIYVLPDLSQYQLSQRDMALVQAMQ 272
QY 339 QGLLESIPLYNQEK 352
DB 273 DGLLP--PLFQPK 284
RESULT 4
P83535
phenylalanine-4-hydroxylase PA0872 [imported] - Pseudomonas aeruginosa (strain PA01)
C:Species: Pseudomonas aeruginosa
C:Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 31-Dec-2000
C:Accession: P83535
R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.; Broman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Lazbig, K.; Lim, J.; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic pathogen.
A:Reference number: A82950; MUID:20437337; PMID:10984043
A:Accession: P83535
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-262 <STO>
A:CROSS-references: GB:AB004522; GB:AB004091; NID:9944768; PIDN:AAG04261.1; GSPDB:GN00142
A:Experimental source: strain PA01
C:Genetics:
A:Gene: phhA; PA0872

Query Match 12.5%; Score 235.5; DB 2; Length 262;
Best Local Similarity 24.5%; Pred. No. 1.4e-11;
Matches 61; Conservative 52; Mismatches 103; Indels 33; Gaps 6;

QY 93 DMVVSTPFFNRLWYRLLSRFSWLKSYCPRFFLDYLEAFGLSDFLDHOAVIKFPELE 152
DB 13 DNGFHYETHEQWNTLITQLKVIQGRACQYLDGIEQLG-----LPHRIQLDIN 67
QY 153 -----THFSYPSVGVAPHOYLSLLQDRYPIASVMTLDKONFSLTPDLIHLGHV 206
DB 68 RVLOATTGWRVARVPALIPFQTFEELASQCFVATFIRTEELDYLDQEPDIFHEIFGHC 127
QY 207 PWLLHPSSEFFINMGRFTKVIKQVALPSKKQRIQTLQSNLIAIVRCFWFTVESGLIE 266
DB 128 PLTLPFWFAEFTHTYGLKGLK-----SKER-----VFLALYNTVEFGLVE 171
QY 267 NHEGRKAYGAVLISSPOELGHAFID-----NVRVLPLELDQIIRLPNTSTPQETILFSIRH 323
DB 172 TDQKKRIYGGGILSPKKTIVYSLSDPLHQAFNPLE---AMRTPYRIDILQPLVFLPDL 228
QY 324 DELVELTSK 332
DB 229 KRLFLQAE 237

RESULT 5

S10489
tryptophan 5-monooxygenase (EC 1.14.16.4) - human
N/Alternate names: tryptophan 5-hydroxylase
C/Species: Homo sapiens (man)
C/Date: 04-Sep-1998 #sequence_revision 04-Sep-1998 #text_change 03-Mar-2000
C/Accession: S10489; S1559
R/Boulard, S.; Darmon, M.C.; Ganem, Y.; Launay, J.M.; Mallet, J.
Nucleic Acids Res. 18, 4257, 1990
A/Title: Complete coding sequence of human tryptophan hydroxylase.
A/Reference number: S10489; MUID:9032431; PMID:2377472
A/Accession: S10489
A/Molecule type: mRNA
A/Residues: 1-444 <BOU>
A/Cross-references: EMBL:X52836; NID:937954; PIDN:CAA37018.1; PID:g37955
R/Tippner, J.P.; Citron, B.A.; Ribeiro, P.; Kaufman, S.
Arch. Biochem. Biophys. 315, 445-453, 1994
A/Title: Cloning and expression of rabbit and human brain tryptophan hydroxylase cDNA in
A/Reference number: S51199; MUID:95077422; PMID:7986090
A/Accession: S51159
A/Status: preliminary
A/Molecule type: mRNA
A/Residues: 1-18, 'T', '20-67', 'T', '69-89', 'TP', '92-96', 'M', '98-99', 'E', '101-103', 'S', '105-150', 'S', '15
35', 'G', '437-444 <TIP>
A/Cross-references: GB:L29306; NID:9531192; PIDN:AAA67050.1; PID:g531193
C/Genetics:
A/Genes: GDB:TPH; TPRH
A/Cross-references: GDB:L20732; OMIM:191060
A/Map position: 11p15.1-11p14.3
C/Function:
A/Description: catalyzes the oxidation of tryptophan to 5'-hydroxytryptophan by tetrahyd
A/Pathway: melatonin biosynthesis; serotonin biosynthesis; tryptophan catabolism
C/Superfamily: phenylalanine 4-monooxygenase
C/Keywords: biotin; iron; melatonin biosynthesis; metalloprotein; monooxygenase; oxid
F:58/Binding site: phosphate (Ser) (covalent) (by cAMP-dependent kinase) #status predict
F:260,443/Binding site: phosphate (Ser) (covalent) (by calmodulin-dependent kinase) #sta
F:272,277,317/Binding site: iron (His, His, Glu) #status predicted

Query Match 12.1%; Score 231.5; DB 1; Length 444;
Best Local Similarity 27.8%; Pred. No. 5.9e-11;
Matches 64; Conservative 43; Mismatches 96; Indels 27; Gaps 4;
QY 107 WYRLSSRFLSKWYCPYRFLDYLEAFGLSDFLDH-----QAVIKFLETHFSY 158
DB 174 WGTVPQELNKLYPTHACR---EYLKNPLLSKYCYGVREDNIPQLEDVSNFLKERTGFSIR 230
QY 159 PVSGFVAPHVYLSLQDRYFPFIASVMTLKDKNFSLTPDLIHDLLGHVPLLPSPSEFF 218
DB 231 PVAGYLSRDLGLAFRVFCHTQYVRHSSDPFYTPPEPTCHELLGHVPLLPSPSAQFS 290
QY 219 INMGLRFTKVKVQALPSKKQRIOTLOSNIATVRCFWFTVSGLTENHEGRKAYGAVL 278
DB 291 QEIG-----LASLGASEAVQKLAT-----CYFTVEFGLCKQDGLRVFGAGL 334
QY 279 ISSPQELGHAFIDNVRVLPDLQIIRLPNTSTPQETLFSIRHDELVE 328
DB 335 LSSISSELKHALSGHAKVFPDKITVKQECILITTFQDVYFVSESFEDAKE 384

RESULT 6
A53452
phenylalanine hydroxylase - Pseudomonas aeruginosa
C/Species: Pseudomonas aeruginosa
C/Date: 26-Jul-1996 #sequence_revision 26-Jul-1996 #text_change 08-Oct-1999
C/Accession: A53452
R/Zhao, G.S.; Xia, T.; Song, J.; Roy, R.A.
Proc. Natl. Acad. Sci. U.S.A. 91, 1366-1370, 1994
A/Title: Pseudomonas aeruginosa possesses homologues of mammalian phenylalanine hydroxyl
A/Reference number: A53452; MUID:94151331; PMID:8108417
A/Accession: A53452
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-262 <RES>
A/Cross-references: GB:M88627; NID:g476740; PIDN:AAA25936.1; PID:g476741

Query Match 12.1%; Score 229.5; DB 2; Length 262;
Best Local Similarity 24.1%; Pred. No. 4.2e-11;
Matches 60; Conservative 52; Mismatches 104; Indels 33; Gaps 6;
QY 93 DMVVSTPFFNRNLYRLLSRFLSKWYCPYRFLDYLEAFGLSDFLDQAVIKFLE 152
DB 13 DNGFTHYPETEHQVWNTLITRQLKVIKGRACQYLDGIEQLG-----LPHRIPQDLDEIN 67
QY 153 -----THFSYVPYSGVAPHOYLSLQDRYFPFIASVMTLKDKNFSLTPDLIHDLLGHV 206
DB 58 RVLQATTCGRVARVPALIPFQTFELLASQCFVATIRTPBELDYLOEPIFHEIFGHC 127
QY 207 PWLLHPSSEFFINMGLRFTKVKVQALPSKKQRIOTLOSNIATVRCFWFTVSGLTIE 266
DB 128 PLLTNPMIAEFTHTYKGLKLA-----SKEER-----VFLARLYWMTTEFGGLVE 171
QY 267 NHEGRKAYGAVLISPPQELGHAFID--NVRVLPDLQIIRLPNTSTPQETLFSIRHF 323
DB 172 TDQGRKIYGGGLISPKETVYSLSDLEPLHQAFNLE---AMRTFYRIDILQPLVFLPDL 228
QY 324 DELVELTSK 332
DB 229 KSLFOLAQE 237

RESULT 7
S51199
tryptophan 5-monooxygenase (EC 1.14.16.4) - rabbit
C/Species: Oryctolagus cuniculus (domestic rabbit)
C/Date: 15-Jul-1995 #sequence_revision 19-Apr-1996 #text_change 31-Mar-2000
C/Accession: S51199; A32699
R/Tippner, J.P.; Citron, B.A.; Ribeiro, P.; Kaufman, S.
Arch. Biochem. Biophys. 315, 445-453, 1994
A/Title: Cloning and expression of rabbit and human brain tryptophan hydroxylase cDNA i
A/Reference number: S51199; MUID:95077422; PMID:7986090
A/Accession: S51199
A/Status: preliminary
A/Molecule type: mRNA
A/Residues: 1-444 <TIP>
A/Cross-references: GB:L29305; NID:9531213; PIDN:AAA67051.1; PID:g531214
R/Grenett, H.E.; Ledley, F.D.; Reed, L.L.; Woo, S.L.C.
Proc. Natl. Acad. Sci. U.S.A. 84, 5530-5534, 1987
A/Title: Full-length cDNA for rabbit tryptophan hydroxylase: functional domains and evo
A/Reference number: A32699; MUID:87289638; PMID:3475690
A/Accession: A32699
A/Status: preliminary
A/Molecule type: mRNA
A/Residues: 1-101, 'L', '103-150', 'L', '152-201', 'ND', '204-206', 'R', '208-389', 'K', '391-444 <GRE>
A/Cross-references: GB:M17250; NID:9165771; PIDN:AAA311487.1; PID:g165772
C/Superfamily: phenylalanine 4-monooxygenase
C/Keywords: biotin; iron; metalloprotein; oxidoreductase; phosphoprotein
F:272,277,317/Binding site: iron (His, His, Glu) #status predicted

Query Match 12.1%; Score 229.5; DB 2; Length 444;
Best Local Similarity 27.4%; Pred. No. 8.5e-11;
Matches 63; Conservative 43; Mismatches 97; Indels 27; Gaps 4;
QY 107 WYRLSSRFLSKWYCPYRFLDYLEAFGLSDFLDHQA-----VTKFLETHFSY 158
DB 174 WGTVPQELNKLYPTHACR---EYLKNPLLSKYCYGVREDNIPQLEDVSNFLKERTGFSIR 230
QY 159 PVSGFVAPHVYLSLQDRYFPFIASVMTLKDKNFSLTPDLIHDLLGHVPLLPSPSEFF 218
DB 231 PVAGYLSRDLGLAFRVFCHTQYVRHSSDPFYTPPEPTCHELLGHVPLLPSPSAQFS 290
QY 219 INMGLRFTKVKVQALPSKKQRIOTLOSNIATVRCFWFTVSGLTENHEGRKAYGAVL 278
DB 291 QEIG-----LASLGASEAVQKLAT-----CYFTVEFGLCKQDGLRVFGAGL 334
QY 279 ISSPQELGHAFIDNVRVLPDLQIIRLPNTSTPQETLFSIRHDELVE 328
DB 335 LSSISSELKHALSGHAKVFPDKITVKQECILITTFQDVYFVSESFEDAKE 384

RESULT 8

A28582
N: tyrosine 3-monooxygenase (EC 1.14.16.2) - quail
N: Alternate names: tyrosine 3-hydroxylase
C: Species: Phasianidae gen. sp. (quail)
C: Date: 19-Nov-1988 #sequence_revision 19-Nov-1988 #text_change 31-Mar-2000
A: Accession: A28582; PH1524
R: Fauquet, M.; Grima, B.; Lancuoux, A.; Mallet, J.
J. Neurochem. 50, 142-148, 1988
A: Title: Cloning of quail tyrosine hydroxylase: amino acid homology with other hydroxylase
A: Reference number: A28582; MUID: 88089590; PMID: 2447231
A: Accession: A28582
A: Molecule type: mRNA
A: Residues: 1-491 <FAU>
A: Cross-references: GB:M24778; NID: g2113649; PIDN: AAA49514.1; PID: g2113650
R: Fauquet, M.; Boni, C.

J. Neurochem. 60, 274-281, 1993
A: Title: The quail tyrosine hydroxylase gene promoter contains an active cyclic AMP-response element
A: Reference number: PH1524; MUID: 93107923; PMID: 80933261
A: Accession: PH1524
A: Molecule type: DNA
A: Residues: 1-30 <FA2>
C: Superfamily: phenylalanine 4-monooxygenase
C: Keywords: bioterrorin; catecholamine biosynthesis; iron; metalloprotein; oxidoreductase;
F: 324,329,369/Binding site: iron (His, His, Glu) #status predicted

Query Match 12.1%; Score 229.5; DB 2; Length 491;
Best Local Similarity 29.1%; Pred. No. 9.7e-11;
Matches 66; Conservative 34; Mismatches 100; Indels 27; Gaps 4;
QY 107 WYLLSSRSLWKSYPFRFLDYLEAFGLSDFLDH-----QAVIKFPELETHFSY 158
DB 226 WKEVYSLKSLYTHACK---EYLRNLLSKYCYREDNVQPLEDVSNFLKERTGFSIR 282
QY 159 PVSGFVAPHOYLSLQDRYFPFIASVMTLDKDNFSLTPDLIHLGHVPMWLLHPSSEFF 218
DB 283 PVGGLSARDPLASLAFVFCQYVHSSPMHSEPCCHELLGHVPMWLLADKTFAPQS 342
QY 219 INNGRLFTKVIKQVLPKQKRIQTLQSNLIAIVRCFWFTVSGLIENHGRKAYGAVL 278
DB 343 QDIG-----LASLGATDEIEKATL-----YMFTEVFGLCRQNGIVKAYGAGL 386
QY 279 ISSPQELGHAFIDNVRVLPLELDQIIRLPNTSTPQETLFSIRHPE 325
DB 387 LSSYGELHSLDEPEVDFDPAAAVQPCDQYQVYFVSESFSD 433

RESULT 9

WHRTW
N: tyrosine 5-monooxygenase (EC 1.14.16.4) - rat
N: Alternate names: tyrosine 5-hydroxylase
C: Species: Rattus norvegicus (Norway rat)
C: Date: 30-Jun-1989 #sequence_revision 30-Jun-1989 #text_change 03-Mar-2000
A: Accession: J00034; A60034; A24367
R: Darmon, M.C.; Guilbert, B.; Leviel, V.; Ehret, M.; Maitre, M.; Mallet, J.
J. Neurochem. 51, 312-316, 1988
A: Title: Sequence of two mRNAs encoding active rat tyrosine hydroxylase.
A: Reference number: J00034; MUID: 88244702; PMID: 3379411
A: Accession: J00034
A: Molecule type: mRNA
A: Residues: 1-444 <DAR>
A: Cross-references: GB:X53501; NID: g57760; PIDN: CAA37579.1; PID: g57761
R: Kim, K.S.; Wessel, T.C.; Stone, D.M.; Carver, C.H.; Joh, T.H.; Park, D.H.
Brain Res. Mol. Brain Res. 9, 277-283, 1991
A: Title: Molecular cloning and characterization of cDNA encoding tyrosine hydroxylase
A: Reference number: A60034; MUID: 91245924; PMID: 1645430
A: Accession: A60034
A: Status: not compared with conceptual translation
A: Molecule type: mRNA
A: Residues: 1-444 <KIM>

A: Experimental source: dorsal raphe nucleus
R: Darmon, M.C.; Grima, B.; Cash, C.D.; Maitre, M.; Mallet, J.

FEBS Lett. 206, 43-46, 1986
A: Title: Isolation of a rat pineal gland cDNA clone homologous to tyrosine and phenylalanine hydroxylase
A: Reference number: A24367; MUID: 87005247; PMID: 2875901
A: Accession: A24367
A: Molecule type: mRNA
A: Residues: 167-261 <DA2>
A: Cross-references: GB:M28000; NID: g207432; PIDN: AAA42262.1; PID: g207433
C: Comment: This enzyme has different physical properties in pineal gland and in dorsal raphe nucleus
C: Function:
A: Description: catalyzes the oxidation of tryptophan to 5'-hydroxytryptophan by tetrahydrobiopterin
A: Pathway: melatonin biosynthesis; serotonin biosynthesis; tryptophan catabolism
C: Superfamily: phenylalanine 4-monooxygenase
C: Keywords: bioterrorin; iron; melatonin biosynthesis; metalloprotein; monooxygenase; oxidoreductase;
F: 58/Binding site: phosphate (Ser) (covalent) (by cAMP-dependent kinase) #status predicted
F: 260,443/Binding site: phosphate (Ser) (covalent) (by calmodulin-dependent kinase) #status predicted
F: 272,277,317/Binding site: iron (His, His, Glu) #status predicted

Query Match 12.0%; Score 227.5; DB 1; Length 444;
Best Local Similarity 27.4%; Pred. No. 1.2e-10;
Matches 63; Conservative 43; Mismatches 97; Indels 27; Gaps 4;
QY 107 WYLLSSRSLWKSYPFRFLDYLEAFGLSDFLDH-----QAVIKFPELETHFSY 158
DB 174 WGTIFRELKSLYTHACK---EYLRNLLSKYCYREDNVQPLEDVSNFLKERTGFSIR 230
QY 159 PVSGFVAPHOYLSLQDRYFPFIASVMTLDKDNFSLTPDLIHLGHVPMWLLHPSSEFF 218
DB 231 PVAGYLSRDFLSGLAFRVFCTQYVHSSDPLYPDPDTCHELLGHVPMWLLAEPSPAQFS 290
QY 219 INNGRLFTKVIKQVLPKQKRIQTLQSNLIAIVRCFWFTVSGLIENHGRKAYGAVL 278
DB 291 QEIG-----LASLGASEETVQKAT-----CYPFTVEFGLCQDQQLRVFAGL 334
QY 279 ISSPQELGHAFIDNVRVLPLELDQIIRLPNTSTPQETLFSIRHPELVE 328
DB 335 LSSISELHSLGSHAKYKVPDPKQCECLITSFQDVFVSESFSDAKE 384

RESULT 10

A24582
N: tyrosine 5-monooxygenase (EC 1.14.16.4) - mouse
C: Species: Mus musculus (house mouse)
C: Date: 22-Jun-1990 #sequence_revision 09-Oct-1992 #text_change 31-Mar-2000
A: Accession: A34582
R: Stoll, J.; Kozak, C.A.; Goldman, D.
Genomics 7, 88-96, 1990
A: Title: Characterization and chromosomal mapping of a cDNA encoding tryptophan hydroxylase
A: Reference number: A34582; MUID: 90243261; PMID: 2110547
A: Accession: A34582
A: Status: preliminary
A: Molecule type: mRNA
A: Residues: 1-447 <STO>
A: Cross-references: GB:J04758; NID: g202113; PIDN: AAA63401.1; PID: g202114
A: Note: the authors translated the codon AAC for residue 405 as Gln
C: Superfamily: phenylalanine 4-monooxygenase
C: Keywords: bioterrorin; iron; metalloprotein; oxidoreductase; phosphoprotein
F: 275,280,320/Binding site: iron (His, His, Glu) #status predicted

Query Match 12.0%; Score 226.5; DB 2; Length 447;
Best Local Similarity 27.4%; Pred. No. 1.5e-10;
Matches 63; Conservative 43; Mismatches 97; Indels 27; Gaps 4;
QY 107 WYLLSSRSLWKSYPFRFLDYLEAFGLSDFLDH-----QAVIKFPELETHFSY 158
DB 177 WGTIFRELKSLYTHACK---EYLRNLLSKYCYREDNVQPLEDVSNFLKERTGFSIR 233
QY 159 PVSGFVAPHOYLSLQDRYFPFIASVMTLDKDNFSLTPDLIHLGHVPMWLLHPSSEFF 218
DB 234 PVAGYLSRDFLSGLAFRVFCTQYVHSSDPLYPDPDTCHELLGHVPMWLLAEPSPAQFS 293
QY 219 INNGRLFTKVIKQVLPKQKRIQTLQSNLIAIVRCFWFTVSGLIENHGRKAYGAVL 278

Db 294 QEIG-----LASIGASSETVOKLAT-----CYFFTVFGLCKQDQQLRVFGAGL 337
QY 279 ISSPOELGHAFIDNVRLVPLELDQIIRLPNTSTPQETLFSIRHFSFSEFF 328
Db 338 LSSISELGHSLSGHAKVFPFKIACKQECUITSFQDVFVSESFEDAKE 387
RESULT 11
JN0068
tyrosine 3-monooxygenase (EC 1.14.16.2) - mouse
N/Alternate names: tyrosine hydroxylase
C/Species: Mus musculus (house mouse)
C/Date: 31-Dec-1991 #sequence_revision 31-Dec-1991 #text_change 31-Mar-2000
A/Title: Primary structure of mouse tyrosine hydroxylase deduced from its cDNA.
A/Reference number: JN0068; MUID:91248263; PMID:1674869
A/Accession: JN0068
R/Chikawa, S.; Sasaki, T.; Negatsu, T.
Biochem. Biophys. Res. Commun. 176, 1610-1616, 1991
A/Title: Primary structure of mouse tyrosine hydroxylase deduced from its cDNA.
A/Reference number: JN0068; MUID:91248263; PMID:1674869
A/Accession: JN0068
A/Molecule type: mRNA
A/Residues: 1-498 <ICH>
A/Cross-references: GB:M9200; NID:G201997; PIDN:AAA40434.1; PID:G201998
A/Experimental source: brain
R/Morgan, W.W.; Bermudez, J.; Sharp, Z.D.
A/Description: DC-12 Nuclear Extracts Produce Tissue-Specific Protection of Several Sequences in the Presence of Tyrosine Hydroxylase
A/Reference number: S21322
A/Accession: S21322
A/Molecule type: DNA
A/Residues: 1-30 <MOR>
A/Cross-references: EMBL:X59503; NID:G55055; PIDN:CAA37580.1; PID:G55056
C/Comment: This enzyme, which requires ferrous iron, catalyzes the hydroxylation of tyrosine in the physiology of adrenergic neurons.
C/Species: Rattus norvegicus (Norway rat)
C/Suprafamily: phenylalanine 4-monooxygenase
C/Keywords: bipterin; iron; metalloprotein; monooxygenase; oxidoreductase; phosphoprotein
F/333,336,376/Binding site: iron (His, His, Glu) #status predicted
Query Match 11.9%; Score 225.5; DB 2; Length 496;
Best Local Similarity 27.5%; Pred. No. 2.1e-10;
Matches 66; Conservative 42; Mismatches 101; Indels 31; Gaps 5;
QY 107 WYRLSSRFLSKSCPRFFLDYLEAFGLSDFLDH-----QAVIKFPELETHFSY 158
Db 233 WKVETATLKGLYATHACR--EHLFAFOLLERYCGYRDSIPQEDVSHFKETGQLR 289
QY 159 PVSGFVAPHQVLSLQDRYFFPIASVWRTLDKDNFSLTPDLIHLGLHVPMLHPSFSEFF 218
Db 290 PVAGLSARDFLASLAFRVFQCTQYIRHASSPMHSPEDCCHELLGHVPMADRTPAQS 349
QY 219 INMGLFTKVKVQALPSKQRIOTLOSNIATVRCFWFTVSGSLIENHGRKAYGAVL 278
Db 350 QDIG-----LASLGASDEEIEKLT-----VYVFTVFEGLCKQNGELKAYGAGL 393
QY 279 ISSPOELGHAFIDNVRLVPLELDQIIRLPNTSTPQETLFSIRHFSFSEFF 334
Db 394 LSSYGLLSLSEEEVFAFDPTAAVQYQDQTVQYVYVSESFSDAKOKLRNVASRIQ 453
RESULT 12
I51567
tryptophan 5-monooxygenase (EC 1.14.16.4) - African clawed frog
C/Species: Xenopus laevis (African clawed frog)
C/Date: 29-May-1998 #sequence_revision 29-May-1998 #text_change 31-Mar-2000
A/Accession: I51567
R/Green, C.B.; Besharse, J.C.
J. Neurochem. 62, 2420-2428, 1994
A/Title: Tryptophan hydroxylase expression is regulated by a circadian clock in Xenopus laevis
A/Reference number: I51567; MUID:94246419; PMID:8189245
A/Accession: I51567
A/Status: Preliminary; translated from GB/EMBL/DBJ
A/Molecule type: mRNA
A/Residues: 1-481 <GRE>

A/Cross-references: GB:L20679; NID:G450644; PIDN:AAA21306.1; PID:G450645
C/Suprafamily: phenylalanine 4-monooxygenase
C/Keywords: bipterin; iron; metalloprotein; oxidoreductase
F/309,314/Binding site: iron (His, His, Glu) #status predicted
Query Match 11.8%; Score 223.5; DB 2; Length 481;
Best Local Similarity 29.5%; Pred. No. 2.9e-10;
Matches 70; Conservative 38; Mismatches 90; Indels 39; Gaps 9;
QY 107 WYRLSSRFLSKSCPRFFLDYLEAFGLSDFLDH-----QAVIKFPELETHFSY 158
Db 211 WGTVPFRELKLYPTHACR---EYLNKLLSKHSCYREDNIPQLEDVSRFLERTGFTIR 267
QY 159 PVSGFVAPHQVLSLQDRYFFPIASVWRTLDKDNFSLTPDLIHLGLHVPMLHPSFSEFF 217
Db 268 PVAGYLSRDFLAGLAFRVFHTCTQYVRH-DSPLATPPEPDTCHELLGHVFLAESPFAQF 326
QY 218 FINMGLFTKVKVQALPSKQRIOTLOSNIATVRCFWFTVSGSLIENHGRKAYGAV 277
Db 327 SQEIG-----LASLGASDEAVQKLT-----CYFFTVFEGLCKQNGELKAYGAG 370
QY 278 LSSPOELGHAFIDNVRLVPLELDQIIRLPNTSTPQETLFSIRHFSFSEFF 334
Db 371 LSSISELGHSLSGHAKVFPD-----PMWTCN-QECI---ITSFQELYFVSESE 417
RESULT 13
WHRTY
tyrosine 3-monooxygenase (EC 1.14.16.2) - rat
N/Alternate names: tyrosine 3-hydroxylase
C/Species: Rattus norvegicus (Norway rat)
C/Date: 04-Dec-1986 #sequence_revision 04-Dec-1986 #text_change 03-Mar-2000
A/Accession: A00510; A44714; S03026; I58264
R/Grima, B.; Lamouroux, A.; Planot, F.; Faucon Biguet, N.; Mallet, J.
Proc. Natl. Acad. Sci. U.S.A. 82, 617-621, 1985
A/Title: Complete coding sequence of rat tyrosine hydroxylase mRNA.
A/Reference number: A00510; MUID:85113249; PMID:2857492
A/Accession: A00510
A/Molecule type: mRNA
A/Residues: 1-498 <ORI>
A/Cross-references: GB:M10244; NID:G207408; PIDN:AAA42257.1; PID:G207409
R/Campbell, D.G.; Hardie, D.G.; Vulliam, P.R.
J. Biol. Chem. 261, 10489-10492, 1986
A/Title: Identification of four phosphorylation sites in the N-terminal region of tyrosine hydroxylase
A/Reference number: A44714; MUID:86278113; PMID:2874140
A/Accession: A44714
A/Status: preliminary
A/Molecule type: protein
A/Residues: 2-12,16-24,38-47,151-157 <CAM>
R/Bonnesfey, E.; Ferrara, P.; Rohrer, F.; Gros, F.; Thibault, J.
Eur. J. Biochem. 174, 685-690, 1988
A/Title: Role of the N-terminus of rat pheochromocytoma tyrosine hydroxylase in the regulation of its activity
A/Reference number: S03026; MUID:88271342; PMID:2899026
A/Accession: S03026
A/Molecule type: protein
A/Residues: 2-26 <BON>
R/Harrington, C.A.; Lewis, E.J.; Krzemien, D.; Chikaraishi, D.M.
Nucleic Acids Res. 15, 2363-2384, 1987
A/Title: Identification and cell type specificity of the tyrosine hydroxylase gene promoter in the rat
A/Reference number: I58264; MUID:87174758; PMID:2882469
A/Accession: I58264
A/Status: translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-30 <RES>
A/Cross-references: EMBL:X04914; NID:G57355; PIDN:CAA28584.1; PID:G57356
C/Function:
A/Description: catalyzes the 3'-hydroxylation of tyrosine to 3',4'-dihydroxyphenylalanine
A/Pathway: catecholamine biosynthesis
A/Note: this is the rate-limiting step in catecholamine biosynthesis
C/Suprafamily: phenylalanine 4-monooxygenase
C/Keywords: bipterin; catecholamine biosynthesis; iron; metalloprotein; monooxygenase;
F/8/Binding site: phosphate (Ser) (covalent) (by unidentified kinase) #status experimentally confirmed
F/19/Binding site: phosphate (Ser) (covalent) (by calmodulin-dependent kinase) #status experimentally confirmed

```
F:40,153/Binding site: phosphate (Ser) (covalent) (by cAMP-dependent kinase) #status exp
F:331,336,376/Binding site: iron (His, His, Glu) #status predicted

Query Match 11.8%; Score 223; DB 1; Length 498;
Best Local Similarity 26.7%; Pred. No. 3.3e-10;
Matches 65; Conservative 39; Mismatches 103; Indels 36; Gaps 5;

QY 112 SRFSLWKS-----YCPFFLDYLEAFGLLSDFLDH-----QAVIKFFLETHF 155
DB 227 AEBIATWKEVYVTLKGLYATHACREHLEGFLERYCGYREDSPQLEDVSVFLKERTGF 286
QY 156 SYVPVSGFVAPHQYLLQDRYFPIASVMRTLDKNFSLTPDLIHDLGHVFWLLHPSPFS 215
DB 287 QLRPVAGLSARDFLASLAFVFOCTQYIRHASSPMHSPPECCHELLGHVFWMLADRTFA 346
QY 216 EFFINMGRLETKVIEKQVALPSKKQRIQTLQSNLIAIVCFWTFVSGLIENHGRKAYG 275
DB 347 QPSQDIG-----LASLGASDEIEKLST-----LYWFTVEFGCKQNGELKAYG 390
QY 276 AVLTSSQDELGHAFIDNVRVLPLELDQIIRLPNTSTPQETLFSIRHF-----DELVELTS 331
DB 391 AGLSSVGEHLHLSSEPEIRAFDPDAAVQPYQDQYQPVYFVSEFSFNDKDKLKNYAS 450
QY 332 KLE 334
DB 451 RIQ 453

RESULT 14
145983
tyrosine 3-monooxygenase (EC 1.14.16.2) - bovine
C:Species: Bos primigenius taurus (cattle)
C:Date: 16-Aug-1996 #sequence_revision 16-Aug-1996 #text_change 31-Mar-2000
C:Accession: J45983
R:D'Mello, S.R.; Weisberg, E.P.; Stachowiak, M.K.; Turzai, L.M.; Gioio, A.E.; Kaplan, B.
J. Neurosci. Res. 19, 440-449, 1988
A:Title: Isolation and nucleotide sequence of a cDNA clone encoding bovine adrenal tyros
A:Reference number: J45983; MUID:88259287; PMID:2898537
A:Accession: J45983
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-491 <DKM>
A:Cross-references: GB:M36784; NID:G163750; PIDN:AAA30779.1; PID:G163751
C:Superfamily: phenylalanine 4-monooxygenase
C:Keywords: biotin; iron; metalloprotein; monooxygenase; oxidoreductase
F:324,329,369/Binding site: iron (His, His, Glu) #status predicted

Query Match 11.7%; Score 221.5; DB 2; Length 491;
Best Local Similarity 27.1%; Pred. No. 4.3e-10;
Matches 65; Conservative 42; Mismatches 102; Indels 31; Gaps 5;

QY 107 WYRLSSRFSLWKSYPFRFFLDYLEAFGLLSDFLDH-----QAVIKFFLETHFSY 158
DB 226 WKEVYVTLRGLYTHACR--EHLAEFELLERFCGYREDRIPOLEDVSVFLKERTGFQLR 282
QY 159 PVSGFVAPHQYLLQDRYFPIASVMRTLDKNFSLTPDLIHDLGHVFWLLHPSPSEFF 218
DB 283 PAAGLLSARDFLASLAFVFOCTQYIRHASSPMHSPPECCHELLGHVFWMLADRTFAQFS 342
QY 219 INMGRLETKVIEKQVALPSKKQRIQTLQSNLIAIVCFWTFVSGLIENHGRKAYGAVL 278
DB 343 QDIG-----LASLGASDEIEKLST-----LYWFTVEFGCKQNGEYKAYGAGL 386
QY 279 ISSQDELGHAFIDNVRVLPLELDQIIRLPNTSTPQETLFSIRHF-----DELVELTSKLE 334
DB 387 LSSVGEHLHLSSEPEIRAFDPDAAVQPYQDQYQPVYFVSEFSFNDKDKLKNYASRIQ 446

RESULT 15
WHU74
tyrosine 3-monooxygenase (EC 1.14.16.2), splice form 4 - human
N:Alternate names: tyrosine 3-hydroxylase
N:Contains: tyrosine 3-monooxygenase, splice form 1; tyrosine 3-monooxygenase, splice fd
```

```
C:Species: Homo sapiens (man)
C:Date: 31-Mar-1989 #sequence_revision 31-Mar-1989 #text_change 16-Jun-2000
C:Accession: A30002; A28825; A60201; JE0012; JE0014; A27791; E27791; C27791; PN:
R:Nagatsu, T.
submitted to GenBank, December 1987
A:Reference number: A94509
A:Accession: A30002
A:Molecule type: mRNA
A:Residues: 1-528 <NAG1>
A:Cross-references: GB:M17589; NID:G339680; PIDN:AAA61179.1; PID:G339681
R:Kaneda, N.; Kobayashi, K.; Ichinose, H.; Kishi, F.; Nakazawa, A.; Kurosawa, Y.; Fujita
Biochem. Biophys. Res. Commun. 146, 971-975, 1987
A:Title: Isolation of a novel cDNA clone for human tyrosine hydroxylase: alternative RN
A:Reference number: A90136; MUID:87298614; PMID:2887169
A:Accession: A26825
A:Molecule type: mRNA
A:Residues: 1-94 <NAG2>
A:Cross-references: GB:M17589; NID:G339680; PIDN:AAA61179.1; PID:G339681
R:Le Bourdellies, B.; Boulard, S.; Boni, C.; Horellou, P.; Dumas, S.; Grima, B.; Mallet
J. Neurochem. 50, 988-991, 1988
A:Title: Analysis of the 5' region of the human tyrosine hydroxylase gene: combinatorial
A:Reference number: A60201; MUID:88117543; PMID:2892893
A:Accession: A60201
A:Molecule type: mRNA
A:Residues: 1-65 <LEB>
A:Cross-references: GB:M24790; NID:G556223; PIDN:AAA61174.1; PID:G556224
R:Kobayashi, K.; Kaneda, N.; Ichinose, H.; Kishi, F.; Nakazawa, A.; Kurosawa, Y.; Fujita
J. Biochem. 103, 907-912, 1988
A:Title: Structure of the human tyrosine hydroxylase gene: alternative splicing from a
A:Reference number: JE0012; MUID:89008200; PMID:2902075
A:Accession: JE0012
A:Molecule type: DNA
A:Residues: 1-30,62-135 <KOB1>
A:Cross-references: GB:D00269; NID:G220099; PIDN:BAA25094.1; PID:G2951764
A:Experimental source: splice form 1
A:Note: this splice form is produced by an alternative donor site within exon 1
A:Accession: JE0013
A:Molecule type: DNA
A:Residues: 1-34,62-135 <KOB2>
A:Cross-references: GB:D00269; NID:G220099; PIDN:BAA25097.1; PID:G2951767
A:Experimental source: splice form 2
A:Accession: JE0014
A:Molecule type: DNA
A:Residues: 1-30,35-135 <KOB3>
A:Cross-references: GB:D00269; NID:G220099; PIDN:BAA25095.1; PID:G2951765
A:Experimental source: splice form 3
A:Note: this splice form is produced by an alternative donor site within exon 1
R:Grima, B.; Lamouroux, A.; Boni, C.; Julien, J.F.; Javoy-Agid, F.; Mallet, J.
Nature 326, 707-711, 1987
A:Title: A single human gene encoding multiple tyrosine hydroxylases with different pre
A:Reference number: A93393; MUID:87173064; PMID:2882428
A:Accession: A27791
A:Molecule type: mRNA
A:Residues: 1-30,62-528 <GRI1>
A:Cross-references: GB:X05290; NID:G32501; PIDN:CAA28908.1; PID:G32502
A:Experimental source: splice form 1
A:Note: this splice form is produced by an alternative donor site within exon 1
A:Accession: B27791
A:Molecule type: mRNA
A:Residues: 1-34,62-528 <GRI2>
A:Cross-references: GB:X05290; NID:G32501
A:Experimental source: splice form 2
A:Accession: C27791
A:Molecule type: mRNA
A:Residues: 30,35-528 <GRI3>
A:Cross-references: GB:X05290; NID:G32501
A:Experimental source: splice form 3
A:Note: this isoform is produced by use of an alternative donor site within exon 1
R:Ichinose, H.; Ohye, T.; Fujita, K.; Yoshida, M.; Ueda, S.; Nagatsu, I.
Biochem. Biophys. Res. Commun. 195, 158-165, 1993
A:Title: Increased heterogeneity of tyrosine hydroxylase in humans.
A:Reference number: PN0575; MUID:93371398; PMID:7689834
A:Accession: PN0575
```


A>Status: translation not shown
A:Molecule type: DNA
A:Residues: 19-30 <ICH1>
A:Accession: FN0582
A>Status: translation not shown
A:Molecule type: DNA
A:Residues: 35-61 <ICH2>
A:Accession: FN0588
A>Status: translation not shown
A:Molecule type: DNA
A:Residues: 62-106 <ICH3>
R:O'Malley, K.L.; Anhalt, M.J.; Martin, B.M.; Kelse, J.R.; Winfield, S.L.; Ginns, E.I.
Biochemistry 26, 2910-2914, 1987
A:Title: Isolation and characterization of the human tyrosine hydroxylase gene: identification of a full-length cDNA clone encoding human tyrosine hydroxylase type
A:Reference number: 152396; MUID:88107612; PMID:2892528
A:Accession: 152396
A>Status: translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-61 <OMA>
A:Cross-references: GB:M18116; NID:G339633; PIDN:AAA77649.1; PID:G1004335
R:Kobayashi, K.; Kaneda, N.; Ichinose, H.; Kishi, F.; Nakazawa, A.; Kurosawa, Y.; Fujita
Nucleic Acids Res. 15, 6733-6737, 1987
A:Title: Isolation of a full-length cDNA clone encoding human tyrosine hydroxylase type
A:Reference number: 138340; MUID:87316931; PMID:2888085
A:Accession: 138340
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-30,35-528 <Y0414>
A:Cross-references: EMBL:Y00414; NID:G37126; PIDN:CAA68472.1; PID:G37127
R:Ginns, E.I.; Rehani, M.; Martin, B.M.; Weller, M.; O'Malley, K.L.; LaMarca, M.E.; McAL
J. Biol. Chem. 263, 7406-7410, 1988
A:Title: Expression of human tyrosine hydroxylase cDNA in invertebrate cells using a bac
A:Reference number: 155282; MUID:88213428; PMID:2896667
A:Accession: 155282
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-30,62-64 <GIN1>
A:Cross-references: GB:M20911; NID:G339636; PIDN:AAA61167.1; PID:G339637
A:Accession: 170056
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-34,62-64 <GIN2>
A:Cross-references: GB:M20912; NID:G339642; PIDN:AAA61168.1; PID:G339643
C:Comment: the expression of the four distinct proteins produced by alternate splicing v
C:Genetics:
A:Gene: GDB:TH
A:Cross-references: GDB:119612; OMIM:191290
A:Map position: 11p15.5-11p15.5
A:Introns: 34/3; 61/3; 135/3
A>Note: the list of introns is incomplete
C:Function:
A>Description: catalyzes the 3'-hydroxylation of tyrosine to 3',4'-dihydroxyphenylalanine
A:Pathway: catecholamine biosynthesis
A>Note: this is the rate-limiting step in catecholamine biosynthesis
C:Superfamily: phenylalanine 4-monooxygenase
C:Keywords: alternative splicing; biotin; catecholamine biosynthesis; iron; metallopro
P1-528/Product: tyrosine 3-monooxygenase, splice form 4 #status predicted <MAT4>
P1-30,35-528/Product: tyrosine 3-monooxygenase, splice form 2 #status predicted <MAT2>
P1-30,62-528/Product: tyrosine 3-monooxygenase, splice form 3 #status predicted <MAT3>
P1-30,62-528/Product: tyrosine 3-monooxygenase, splice form 1 #status predicted <MAT1>
F19/Binding site: phosphate (Thr) (covalent) (by unidentified kinase) #status predicted
F19/Binding site: phosphate (Ser) (covalent) (by calmodulin-dependent kinase) #status p
F71.183/Binding site: phosphate (Ser) (covalent) (by cAMP-dependent kinase) #status pre
F71.183/Binding site: phosphate (His, His, Glu) #status predicted
F71.183/Binding site: iron (His, His, Glu) #status predicted
Query Match 11.6%; Score 220; DB 1; Length 528;
Best Local Similarity 26.7%; Pred. No. 6.2e-10;
Matches 65; Conservative 38; Mismatches 104; Indels 36; Gaps 5;
QY 112 SSRFSLWKS-----YCPRFLLDYLEAFGLSLDFLDH-----QAVIKFFPELETHF 155
DB 257 ABEIATWKEVYTLKGLYATHACGEHLEAFALLERFSGYREDNIPQLEDVSRFLKERTGF 316

QY 156 SYYPVSGFVAFPHOYLKLDQRYFFFIASVMRTLDKDNFSLTDPDLIHLGHVPLHHSFS 215
DB 317 QLRPVAGLLSARDLFLASLAFVFOCTQVIRHASSPMHSPEDCCHELLGHVPLADRTFA 376
QY 216 EFFINMGELFTVKIEKQALPSKKORICTQSLNIAIVRCFWFTVESGLTENHGRKAYG 275
DB 377 QFSQDIG-----LASIGASDEEIEKJSTLS-----WFTVEFGLCKONGEVKAYG 420
QY 276 AVLISSPQELGHAFIDNVRLPDLQIIRLPFNTSTPQETLFSIRHP-----DELVELTS 331
DB 421 AGLSSSYGELLHCLSEEPERAFPEAAAVQPYQDTYQSVYFVSEFSFSDAKDKLRSVAS 480
QY 332 KLE 334
DB 481 RIQ 483

Search completed: March 25, 2004, 14:14:33
Job time : 22 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 25, 2004, 14:10:43 ; Search time 17 Seconds
(without alignments)
1108.788 Million cell updates/sec

Title: US-09-438-185A-1047

Perfect score: 1889

Sequence: 1 VHYCERTLDPKYIKIALKL.....ESIPLYNQBYKLSGFVLCQ 362

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_42.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Length | DB ID | Description |
|------------|-------|--------------|-------|---------------|
| 1 | 1886 | 99.8 | 362 | 1 AAAH_CHLNP |
| 2 | 256.5 | 13.6 | 269 | 1 PH4H_VIBCH |
| 3 | 235.5 | 12.5 | 262 | 1 PH4H_PSEAE |
| 4 | 231 | 12.2 | 445 | 1 TP1H_CHICK |
| 5 | 230.5 | 12.2 | 444 | 1 TP1H_HUMAN |
| 6 | 229.5 | 12.1 | 491 | 1 TY3H_PHASP |
| 7 | 227.5 | 12.0 | 444 | 1 TP1H_RAT |
| 8 | 226.5 | 12.0 | 447 | 1 TP1H_MOUSE |
| 9 | 225.5 | 11.9 | 444 | 1 TP1H_RABIT |
| 10 | 225.5 | 11.9 | 488 | 1 TY3H_ANGAN |
| 11 | 225.5 | 11.9 | 498 | 1 TY3H_MOUSE |
| 12 | 223.5 | 11.8 | 481 | 1 TP3H_XENLA |
| 13 | 223 | 11.8 | 498 | 1 TY3H_RAT |
| 14 | 221.5 | 11.7 | 490 | 1 TY3H_BOVIN |
| 15 | 221 | 11.7 | 275 | 1 PH4H_RHILLO |
| 16 | 220 | 11.6 | 528 | 1 TY3H_HUMAN |
| 17 | 219 | 11.6 | 579 | 1 TY3H_DROME |
| 18 | 216.5 | 11.5 | 465 | 1 TY3H_SCHMA |
| 19 | 216.5 | 11.5 | 490 | 1 PH4H_HUMAN |
| 20 | 214 | 11.3 | 453 | 1 PH4H_RAT |
| 21 | 213.5 | 11.3 | 452 | 1 PH4H_HUMAN |
| 22 | 213.5 | 11.3 | 485 | 1 TP2H_RAT |
| 23 | 212.5 | 11.2 | 453 | 1 PH4H_MOUSE |
| 24 | 205.5 | 10.9 | 488 | 1 TP2H_MOUSE |
| 25 | 201 | 10.6 | 313 | 1 PH4H_RALSO |
| 26 | 195 | 10.3 | 297 | 1 PH4H_CHRVO |
| 27 | 194 | 10.3 | 452 | 1 PH4H_DROME |
| 28 | 189.5 | 10.0 | 294 | 1 PH4H_CAUCR |
| 29 | 178.5 | 9.4 | 524 | 1 TY3H_CABEL |
| 30 | 175.5 | 9.3 | 457 | 1 PH4H_CABEL |
| 31 | 106.5 | 5.6 | 506 | 1 MATK_ARAAL |
| 32 | 106.5 | 5.6 | 1501 | 1 NINC_DROME |
| 33 | 106 | 5.6 | 504 | 1 MATK_CARAN |

| | | | | |
|----|-------|-----|------|--------------|
| 34 | 104.5 | 5.5 | 2331 | 1 RRPL_MABVP |
| 35 | 102 | 5.4 | 563 | 1 YAS8_SCHPO |
| 36 | 100.5 | 5.3 | 1157 | 1 N133_YEAST |
| 37 | 100.5 | 5.3 | 2212 | 1 RRPL_EBOZM |
| 38 | 99.5 | 5.3 | 513 | 1 MATK_SPOIN |
| 39 | 99 | 5.2 | 924 | 1 SECS_RAT |
| 40 | 98.5 | 5.2 | 1024 | 1 Y075_MYCGE |
| 41 | 98 | 5.2 | 924 | 1 SEC5_MOUSE |
| 42 | 97.5 | 5.2 | 504 | 1 MATK_LEPCM |
| 43 | 97 | 5.1 | 492 | 1 CPL1_PANAR |
| 44 | 97 | 5.1 | 1050 | 1 EX5B_CHLNP |
| 45 | 96.5 | 5.1 | 512 | 1 G6PD_CHLNP |

ALIGNMENTS

RESULT 1

| ID | AAAH_CHLNP | STANDARD; | PRT; | 362 AA. |
|----|--|-----------|------|---------|
| AC | Q9Z6L3; | | | |
| DT | 28-FEB-2003 (Rel. 41, Created) | | | |
| DT | 28-FEB-2003 (Rel. 41, Last sequence update) | | | |
| DT | 10-OCT-2003 (Rel. 42, Last annotation update) | | | |
| DE | Probable aromatic amino acid hydroxylase (EC 1.14.16.-). | | | |
| GN | CPN1046 OR CP0806 OR CP01046 OR CPB1086. | | | |
| OS | Chlamydia pneumoniae (Chlamydia pneumoniae) | | | |
| OC | Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydia | | | |
| OX | NCBI_TaxID=83558; | | | |
| RN | [1] | | | |
| RP | SEQUENCE FROM N.A. | | | |
| RC | STRAIN=RW1029. | | | |
| RA | Kalman S., Mitchell W., Marathe R., Lammel C., Fan J., Hyman R.W., | | | |
| RA | Olinger L., Grimwood J., Davis R.W., Stephens R.S.; | | | |
| RT | "Comparative genomes of Chlamydia pneumoniae and C. trachomatis." | | | |
| RL | Nat. Genet. 21:385-389(1999). | | | |
| RN | [2] | | | |
| RP | SEQUENCE FROM N.A. | | | |
| RC | STRAIN=AR39; | | | |
| RA | MEDLINE=20150255; PubMed=10684935; | | | |
| RA | Read T.D., Brunham R.C., Shen C., Gill S.R., Heidelberg J.F., | | | |
| RA | White O., Hickey E.K., Peterson J., Utterback T., Berry K., Bass S., | | | |
| RA | Linher K., Weidman J., Khouri H., Craven B., Bowman C., Dodson R., | | | |
| RA | Gwinn M., Nelson W., DeBoy R., Kolonay J., McClarty G., Salzberg S.L., | | | |
| RA | Eisen J., Fraser C.M.; | | | |
| RT | "Genome sequences of Chlamydia trachomatis MoPn and Chlamydia | | | |
| RT | pneumoniae AR39." | | | |
| RL | Nucleic Acids Res. 28:1397-1406(2000). | | | |
| RN | [3] | | | |
| RP | SEQUENCE FROM N.A. | | | |
| RC | STRAIN=J138; | | | |
| RA | MEDLINE=20330349; PubMed=10871362; | | | |
| RA | Shirai M., Hirakawa H., Kimoto M., Tabuchi M., Kishi F., Ouchi K., | | | |
| RA | Shiba T., Ishii K., Hattori M., Kuhara S., Nakazawa T.; | | | |
| RT | "Comparison of whole genome sequences of Chlamydia pneumoniae J138 | | | |
| RT | from Japan and CWL029 from USA." | | | |
| RL | Nucleic Acids Res. 28:2311-2314(2000). | | | |
| RN | [4] | | | |
| RP | SEQUENCE FROM N.A. | | | |
| RC | STRAIN=TW-183; | | | |
| RA | Geng M.M., Schumacher A., Muehldorfer I., Bensch K.W., Schaefer K.P., | | | |
| RA | Schneider S., Pohl T., Essig A., Marre R., Melchers K.; | | | |
| RT | "The genome sequence of Chlamydia pneumoniae TW183 and comparison with | | | |
| RT | other Chlamydia strains based on whole genome sequence analysis." | | | |
| RL | Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases. | | | |
| CC | !- COPACTOR: Binds 1 ferrous ion (By similarity). | | | |
| CC | !- SIMILARITY: Belongs to the biotin-dependent aromatic amino acid | | | |
| CC | hydroxylase family. | | | |
| CC | This SWISS-PROT entry is copyright. It is produced through a collaboration | | | |
| CC | between the Swiss Institute of Bioinformatics and the EMBL outstation - | | | |
| CC | the European Bioinformatics Institute. There are no restrictions on its | | | |

CG use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

 DR EMBL; A0001685; AAD19183.1; -
 DR EMBL; A0002240; AAF73705.1; -
 DR EMBL; A002548; BAA9253.1; -
 DR EMBL; A0017160; RAP99015.1; -
 DR PIR; C86621; C86621.
 DR PIR; E72002; E72002.
 DR TIGR; CP0806; -
 DR InterPro; IPR001273; Aaa hydroxylase.
 DR Pfam; PF00351; bioprotein H; 1.
 DR PRINTS; PR00372; FWHYDRXLAQE.
 DR ProDom; PD002559; Aaa hydroxylase; 1.
 DR PROSITE; PS00367; BIOTERIN_HYDROXYL; 1.
 KW Oxidoreductase; Monooxygenase; Iron; Complete proteome.
 FT METAL 200 IRON (POTENTIAL).
 FT METAL 205 IRON (POTENTIAL).
 FT CONFLICT 131 131 E -> D (IN REF. 4).
 SQ SEQUENCE 362 AA; 42513 MW; 01B89BB4B4FE593B CRC64;
 Query Match 99.8%; Score 1886; DB 1; Length 362;
 Best Local Similarity 99.7%; Pred. No. 2.3e-140; Indels 0; Gaps 0;
 Matches 361; Conservative 1; Mismatches 0;
 QY 1 VHYCERTLPKYLKIALKRLQSLFFQNSQSLQRAYSTPYSYRIILOKENKEKQALA 60
 DB 1 MHYCERTLPKYLKIALKRLQSLFFQNSQSLQRAYSTPYSYRIILOKENKEKQALA 60
 QY 61 RHKCIISLFFKNLLPVHLLSKNORECSDTMAVSTPFFNRLWLLSSRSLWKS 120
 DB 61 RHKCIISLFFKNLLPVHLLSKNORECSDTMAVSTPFFNRLWLLSSRSLWKS 120
 QY 121 YCPRFFLDVLEAFGLSDFLDQAVIKFPELETHFSYYPVSGFVAPHQVLSLLQDRYFPI 180
 DB 121 YCPRFFLDVLEAFGLSDFLDQAVIKFPELETHFSYYPVSGFVAPHQVLSLLQDRYFPI 180
 QY 181 ASVMTLDKDNFSLTDLIDHLLGHVPWLLHPSFSEFFINMGRLETKVIEKQALPSKKQ 240
 DB 181 ASVMTLDKDNFSLTDLIDHLLGHVPWLLHPSFSEFFINMGRLETKVIEKQALPSKKQ 240
 QY 241 RIOTLQSLNIAIVRCFWFTVESGLIENHGRKAYGAVLSSQELGHAFIDNVRVLPLEL 300
 DB 241 RIOTLQSLNIAIVRCFWFTVESGLIENHGRKAYGAVLSSQELGHAFIDNVRVLPLEL 300
 QY 301 DQIIRLPFNSTPQETLFSIRHFDDELVELTSKLEWMLDQGLLESIPLYNQEKYLSGFEVL 360
 DB 301 DQIIRLPFNSTPQETLFSIRHFDDELVELTSKLEWMLDQGLLESIPLYNQEKYLSGFEVL 360
 QY 361 CQ 362
 DB 361 CQ 362
 RESULT 2
 ID PH4H VIBCH STANDARD; PRT; 289 AA.
 AC QXLB8; 2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Phenylalanine-4-hydroxylase (EC 1.14.16.1) (PAH) (Phe-4-
 DE monooxygenase)
 GN PHA OR VCA0828.
 OS Vibrio cholerae.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
 OC Vibrionaceae; Vibrio.
 OX NCBI_TaxID=666;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=El Tor N16961 / Serotype O1;

 RX MEDLINE=20406833; PubMed=10952301;
 RA Heidelberg J.F., Eisen J.A., Nelson W.C., Clayton R.A., Gwin M.L.,
 RA Dodson R.J., Haft D.H., Hickey E.K., Peterson J.D., Unayam L.A.,
 RA Gill S.R., Nelson K.E., Read T.D., Tettelin H., Richardson D.,
 RA Ermolaeva M.D., Vamathevan J., Bass S., Qin H., Dragoi I., Seilers P.,
 RA McDonald L., Utterback T., Fleischmann R.D., Nierman W.C., White O.,
 RA Salzberg S.L., Smith H.O., Colwell R.R., Mekalanos J.J., Venter J.C.,
 RA Fraser C.M.;
 RT "DNA sequence of both chromosomes of the cholera pathogen Vibrio
 RT cholerae";
 RL Nature 406:477-483(2000).
 CC -!- CARBOLITIC ACTIVITY: L-phenylalanine + tetrahydrobiopterin + O(2) =
 CC L-tyrosine + 4-alpha-hydroxytetrahydrobiopterin.
 CC -!- COFACTOR: Binds 1 ferrous ion (By similarity).
 CC -!- PATHWAY: Catabolism of phenylalanine; first (rate-limiting) step.
 CC -!- SIMILARITY: Belongs to the bioprotein-dependent aromatic amino acid
 CC hydroxylase family.

 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; A0004410; AAF96726.1; -
 DR PIR; D82413; D82413.
 DR HSSP; P04177; ITOH.
 DR TIGR; VCA0828; -
 DR InterPro; IPR001273; Aaa hydroxylase.
 DR InterPro; IPR005960; Phenylalanyl 4-OHaseM.
 DR Pfam; PF00351; bioprotein H; 1.
 DR PRINTS; PR00372; FWHYDRXLAQE.
 DR ProDom; PD002559; Aaa hydroxylase; 1.
 DR TIGRfams; TIGR01267; Phehydrox mono; 1.
 DR PROSITE; PS00367; BIOTERIN_HYDROXYL; 1.
 KW Oxidoreductase; Monooxygenase; Phenylalanine catabolism; Iron;
 KW Complete proteome.
 FT METAL 144 144 IRON (POTENTIAL).
 FT METAL 149 149 IRON (POTENTIAL).
 SQ SEQUENCE 289 AA; 33445 MW; 2D6B831C6E31D521 CRC64;
 Query Match 13.6%; Score 256.5; DB 1; Length 289;
 Best Local Similarity 25.6%; Pred. No. 5.4e-13;
 Matches 65; Conservative 56; Mismatches 108; Indels 25; Gaps 6;
 QY 106 LWYLLSSRSLWKSYPFRFLDYLEAFGLSDFLDQAVI-KFPELETHFSYYPVSGFV 164
 DB 49 VWHELIITRQEVVKTACQAYLDGLANMLNLTDLRLPQLPEINRVLORETGWQVEPVALI 108
 QY 165 APHOYLSLQDRYFPPIASVMRTLDKONFSLTDLIDHLLGHVPWLLHPSFSEFFINMGR 224
 DB 109 SDFRFFALLADKKFPVATFRRREEDYLOEPFFHEVYGHCAMLTHPDPFAATHYVGL 168
 QY 225 FTKVIEKQALPSKKQRIQTLQSNLIAIVRCFWFTVESGLIENHGRKAYGAVLSSQPE 284
 DB 169 GAKATPKERSY-----LRLYFTVEFLVQVGGQTKYGGGILSSPGE 212
 QY 285 LGHAFIDNV-RVLPLELDQIIRLPFNSTPQETLFSIRHFDDELVELTSK---LEWM-LD 338
 DB 213 TLVASESTIPKREPFIDMQLRTPIRIDINQPIYVLPDLSQLYLSQSDVMALVWQAMQ 272
 QY 339 QGLLESIPLYNQEK 352
 DB 273 DGLLP--PLFQPK 284
 RESULT 3
 PH4H PSEAE
 ID PH4H PSEAE STANDARD; PRT; 262 AA.
 AC P43334;
 DT 01-NOV-1995 (Rel. 32, Created)

DT 16-OCT-2001 (Rel. 40, Last sequence update)
DE 10-OCT-2003 (Rel. 42, Last annotation update)
DE Phenylalanine-4-hydroxylase (EC 1.14.16.1) (PAH) (Phe-4-
monooxygenase).
GN PHA OR PA0872.
OS Pseudomonas aeruginosa.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Pseudomonadaceae; Pseudomonas.
OX NCBI_TaxID=287;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 15692 / PA01;
RX MEDLINE=20437337; PubMed=10984043;
RA Stover C.K., Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Warren P.,
RA Hickey M.J., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,
RA Garber R.L., Goltry L., Tolentino E., Westbrook-Hadman S., Yuan Y.,
RA Brody L.L., Coulter S.N., Folger K.R., Kas A., Larbig K., Lim R.M.,
RA Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,
RA Reizer J., Sailer M.H., Hancock R.E.W., Lory S., Olson M.V.;
RA "Complete genome sequence of Pseudomonas aeruginosa PA01, an
RT opportunistic pathogen";
RL Nature 406:959-964(2000).
CC -!- CATALYTIC ACTIVITY: L-phenylalanine + tetrahydrobiopterin + O(2) =
CC L-tyrosine + 4-alpha-hydroxytetrahydrobiopterin.
CC -!- COFACTOR: Binds 1 ferrous ion.
CC -!- PATHWAY: Catabolism of phenylalanine; first (rate-limiting) step.
CC -!- SUBUNIT: Monomer.
CC -!- SIMILARITY: Belongs to the biopterin-dependent aromatic amino acid
CC hydroxylase family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; M8627; AAA25936.1; -.
CC EMBL; AE004522; AAG04261.1; -.
CC PIR; A53452; A53452.
CC PIR; F83535; F83535.
CC DR HSP; P04177; ITOH.
DR InterPro; IPR001273; Aaa_hydroxylase.
DR InterPro; IPR005960; Phenylalanyl-OHase.
DR Pfam; PF0351; biopterin_H; 1.
DR PRINTS; PR00372; FYWHYDRXLAASE.
DR ProDom; PD002559; Aaa_hydroxylase; 1.
DR TIGRFAMs; TIGR01267; Phe4hydrox mono; 1.
DR PROSITE; PS00367; BIOPTERIN HYDROXYL; 1.
KW Oxidoreductase; Monooxygenase; Phenylalanine catabolism; Iron;
FT Complete proteome.
KW METAL 121 121 IRON (BY SIMILARITY).
FT METAL 126 126 IRON (BY SIMILARITY).
FT CONFLICT 135 135 F -> L (IN REF. 1).
SQ SEQUENCE 262 AA; 30322 MW; A5665839C5961AA5 CRC64;
Query Match 12.5%; Score 235.5; DB 1; Length 262;
Best Local Similarity 24.5%; Pred. No. 2.1e-11;
Matches 61; Conservative 52; Mismatches 103; Indels 33; Gaps 6;
QY 93 DNAVSTPFFNNLWYLLSRFSRWKGYCPRFFLDYLEAFGLLSDFLDHQAVIKFFLE 152
DB 13 DNGFIHPETEPHQVNTLITRLKVIEGRACQYLDGIEQLG----LPHRIQLQLEIN 67

QY 153 -----THFSYYPVSGFVAPHOYLSLQDRYFFPIASVMRTLDKONFSLTPDLIHLGLHV 206
DB 68 RVLQATTGWRVARVPALIPFQTFELLASQFPVATEIRTPPELDYLOEPDIFHFQHC 127
QY 207 PMLHPSPSEFFINMGRFLTKVIEKQALPSKKQRIQIQLQSNLIAIVRCFNFVSGGLIE 266
DB 128 PLLTNPWFAEFTHTYKLGKA-----SKER-----VFLARLYWNTIEFGLVE 171
QY 267 NHEGRKAYGAVLISSPQELGHAFID---NVRVLPLELQIIRLPENTSTPQTLFSIRHF 323
DB 172 TDQKRIYGGILSPKVTYLSDEPLHQAHPLE---AMETPYRIDILOPLYFVLPL 228
QY 324 DELVELTSK 332
DB 229 KRLFQLAQE 237
RESULT 4
TPH1_CHICK STANDARD; PRT; 445 AA.
AC P70080;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Tryptophan 5-hydroxylase 1 (EC 1.14.16.4) (Tryptophan 5-
monooxygenase 1).
GN TPH1 OR TPH.
OS Gallus gallus (Chicken).
OC Archaeoptera; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=White Leghorn; TISSUE=Pineal gland;
RX MEDLINE=97072811; PubMed=8915576;
RA Flores J.C., Seidenman K.J., Barrett R.K., Sangoram A.M.,
RA Takahashi J.S.;
RT "Molecular cloning of chick pineal tryptophan hydroxylase and
RT circadian regulation of its mRNA levels";
RL Brain Res. Mol. Brain Res. 42:25-30(1996).
CC -!- CATALYTIC ACTIVITY: L-tryptophan + tetrahydrobiopterin + O(2) = 5-
CC hydroxy-L-tryptophan + 4-alpha-hydroxytetrahydrobiopterin.
CC -!- COFACTOR: Ferrous ion.
CC -!- PATHWAY: Serotonin biosynthesis; first step.
CC -!- PATHWAY: Melatonin biosynthesis; first step.
CC -!- SUBUNIT: Multimer of identical subunits (By similarity).
CC -!- SIMILARITY: Belongs to the biopterin-dependent aromatic amino acid
CC hydroxylase family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; U26428; AAC60036.1; -.
CC HSP; P04176; 1PHZ.
DR InterPro; IPR001273; Aaa_hydroxylase.
DR InterPro; IPR002912; ACT.
DR InterPro; IPR005963; Tyr_5_monox.
DR Pfam; PF01842; ACT; 1.
DR Pfam; PF00351; biopterin_H; 1.
DR PRINTS; PR00372; FYWHYDRXLAASE.
DR ProDom; PD002559; Aaa_hydroxylase; 1.
DR TIGRFAMs; TIGR01270; Trp_5_monox; 1.
DR PROSITE; PS00367; BIOPTERIN HYDROXYL; 1.
KW Oxidoreductase; Monooxygenase; Serotonin biosynthesis; Iron;
KW Phosphorylation.
FT MOD_RES 58 58 PHOSPHORYLATION (BY PKA) (POTENTIAL).

174 WCTVFEELKLYPTHACR---EYLKXNLLSKYGVREDNIQLELDVSNFLKERTGFSIR 230
 159 PVSGFVAPHOYLSLLQDRFPPTASVMTLDKNFSLTLPDLIHLGHVFWLLHPSFSEFF 218
 231 PVAGYLSRDFLSGLAFRVFCHTQYVRHSDDPFYTPEDTCHELLGHVPLLAEPSPAFOS 290
 219 INNGRLFTKVIKQVQALPSKKQRIQTLQSNLIAIVRCFWFTVESGLIENHEGRKAYGAVL 278
 291 QEIG-----LASLGASEAVQKAT-----CYFTVEFGCKDQGLRVFAGL 334
 279 ISSPQELCHAFIDNVRVLPLELDQIIRLPFNSTPQETLFSIRHDELVE 328
 335 LSSISELXGALSGHAKVKKPPDKITCKQECILITTFQDVVFVSSEFSAKE 384

RESULT 6
 TY3H_PHASP STANDARD; PRT; 491 AA.
 AC P11982;
 DT 01-OCT-1989 (Rel. 12, Created)
 DT 01-OCT-1989 (Rel. 12, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Tyrosine 3-monooxygenase (EC 1.14.16.2) (Tyrosine 3-hydroxylase) (TH).
 GN TH.
 OS Phasianidae sp. (Quail).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae.
 OX NCBI_TaxID=9006;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Adrenal gland;
 RX MEDLINE=88089590; PubMed=2447231;
 RA Fauquet M., Grima B., Lamouroux A., Mallet J.;
 RT "Cloning of quail tyrosine hydroxylase: amino acid homology with
 other hydroxylases discloses functional domains."
 RL J. Neurochem. 50:142-148(1986).
 CC - FUNCTION: Plays an important role in the physiology of adrenergic
 CC neurones.
 CC - CATALYTIC ACTIVITY: L-tyrosine + tetrahydropteridine + O(2) = 3,4-
 CC dihydroxy-L-phenylalanine + dihydropteridine + H(2)O.
 CC - COFACTOR: Ferrous ion.
 CC - ENZYME REGULATION: Phosphorylation leads to an increase in the
 CC catalytic activity.
 CC - PATHWAY: Catecholamine biosynthesis; first step.
 CC - SIMILARITY: Belongs to the bipterin-dependent aromatic amino acid
 CC hydroxylase family.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; M24778; AAA49514.1; -;
 CC PIR; A28582; A28582.
 CC HSSP; P04177; 1TCH.
 CC InterPro; IPR001273; Aaa hydroxylase.
 CC InterPro; IPR005962; Tyr_3_monox.
 CC Pfam; PF00351; bipterin H; 1.
 CC PRINTS; PR00372; FYWHYDRXLASE.
 CC ProDom; PD002559; Aaa hydroxylase; 1.
 CC TIGRFAMs; TIGR01269; Tyr_3_monox; 1.
 CC PROSITE; PS00367; BIPTERIN_HYDROXYL; 1.
 CC Catecholamine biosynthesis; Oxidoreductase; Monooxygenase; Iron;
 CC Neurotransmitter biosynthesis; Phosphorylation.
 CC MOD_RES 40 40 PHOSPHORYLATION (BY PKA) (BY SIMILARITY).
 CC DOMAIN 51 55 POLY-ALA.
 CC FT METAL 324 324 IRON (BY SIMILARITY).
 CC FT METAL 329 329 IRON (BY SIMILARITY).
 CC FT METAL 369 369 IRON (BY SIMILARITY).
 CC -SEQUENCE 491 AA; 56066 MW; AFB363220F70C0A0 CRC64;

Query Match 12.1%; Score 229.5; DB 1; Length 491;
 Best Local Similarity 29.1%; Pred. No. 1.4e-10;
 Matches 66; Conservative 34; Mismatches 100; Indels 27; Gaps 4;
 QY 107 WYRLSSRFSWKSVCPRFFLDYLFARGLSDFLDH-----QAVIKFELETHESYY 158
 DB 226 WKYVSTLKSLYPTHACK--EYLEAFNLLEKFCGYNNENNIPOLEEVSRFLKERTGFOLR 282
 QY 159 PVSGFVAPHOYLSLLQDRFPPTASVMTLDKNFSLTLPDLIHLGHVFWLLHPSFSEFF 218
 DB 283 PVRLGSLGARDFLASLAFVFCQTOYIRHASPSPHSPEDCCHELLGHVPLADKTFAPQS 342
 QY 219 INNGRLFTKVIKQVQALPSKKQRIQTLQSNLIAIVRCFWFTVESGLIENHEGRKAYGAVL 278
 DB 343 QDIG-----LASLGATDEIEKLATL-----YFTVEFGLCRQNGIVKAYGAGL 386
 279 ISSPQELCHAFIDNVRVLPLELDQIIRLPFNSTPQETLFSIRHFE 325
 387 LSSYGELIHSLSDEPEVRDFDPAAVQPCDQDPYQPVFVSESFSD 433

RESULT 7
 TPH1_RAT STANDARD; PRT; 444 AA.
 ID TPH1_RAT
 AC P09810;
 DT 01-WAR-1989 (Rel. 10, Created)
 DT 01-OCT-1989 (Rel. 12, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Tryptophan 5-hydroxylase 1 (EC 1.14.16.4) (Tryptophan 5-
 DE monooxygenase 1).
 GN TPH1 OR TPH.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Wistar; TISSUE=Pineal gland;
 RX MEDLINE=88244702; PubMed=3379411;
 RA Darmon M.C., Guibert B., Levial V., Ehret M., Maitre M., Mallet J.;
 RT "Sequence of two mRNAs encoding active rat tryptophan hydroxylase."
 RL J. Neurochem. 51:312-316(1988).
 CC [2]
 CC SEQUENCE FROM N.A.
 CC MEDLINE=91245924; PubMed=1645430;
 CC Kim K.S., Wessel T.C., Stone D.M., Carver C.H., Joh T.H., Park D.H.;
 RT "Molecular cloning and characterization of cDNA encoding tryptophan
 RL hydroxylase from rat central serotonergic neurons."
 RN Brain Res. Mol. Brain Res. 9:277-283(1991).
 CC [3]
 CC SEQUENCE OF 167-261 FROM N.A.
 CC MEDLINE=87005247; PubMed=2675901;
 CC Darmon M.C., Grima B., Cash C.D., Maitre M., Mallet J.;
 RT "Isolation of a rat pineal gland cDNA clone homologous to tyrosine
 RL and phenylalanine hydroxylases."
 CC FEBS Lett. 206:43-46(1986).
 CC - CATALYTIC ACTIVITY: L-tryptophan + tetrahydrobiopterin + O(2) = 5-
 CC hydroxy-L-tryptophan + 4-alpha-hydroxytetrahydrobiopterin.
 CC - COFACTOR: Ferrous ion.
 CC - PATHWAY: Serotonin biosynthesis; first step.
 CC - PATHWAY: Melatonin biosynthesis; first step.
 CC - SUBUNIT: Multimer of identical subunits.
 CC - SIMILARITY: Belongs to the bipterin-dependent aromatic amino acid
 CC hydroxylase family.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

```

CC EMBL; M28000; AAA42262.1; -.
CC EMBL; X53501; CAA37579.1; -.
CC PIR; J04117; ITOH.
CC HSSP; P04117; ITOH.
CC InterPro; IPR001273; Aaa_hydroxylase.
CC InterPro; IPR002912; ACT.
CC InterPro; IPR005963; Tyr_5_monox.
CC Pfam; PF01842; ACT; 1.
CC Pfam; PF00351; bioprotein H; 1.
CC PRINTS; PR00372; FWHYDRXLA.
CC ProDom; PD002559; Aaa_hydroxylase; 1.
CC TIGRFAMs; TIGR01270; Trp_5_monoox; 1.
CC PROSITE; PS00367; BIOPROTEIN HYDROXYL; 1.
CC Oxidoreductase; Monooxygenase; Serotonin biosynthesis; Iron;
KW Phosphorylation.
FT MOD_RES 58 PHOSPHORYLATION (BY PKA) (POTENTIAL).
FT METAL 272 IRON (BY SIMILARITY).
FT METAL 277 IRON (BY SIMILARITY).
FT METAL 317 IRON (BY SIMILARITY).
SQ SEQUENCE 444 AA; 51068 MW; C3CF5245727CC825 CRC64;

Query Match 12.0%; Score 227.5; DB 1; Length 444;
Best Local Similarity 27.4%; Pred. No. 1.7e-10;
Matches 63; Conservative 43; Mismatches 97; Indels 27; Gaps 4;

QY 107 WYRLSSRFLSKYSCPRFFLDYLEAFGLSLDFLDH-----QAVIKFFLETHFSYY 158
DB 174 WGTIFRELNKLPHYTHACR--EYLRLPLLSKYCGYREDNVPQLEDVSNFLKERTGFSIR 230
QY 159 PVSGFVAPHQVLSLQDRYFPIASVMTLDKNFSLTPDLIHLLGHVPLHLPFSSEFF 218
DB 231 PVAGYLSPRDFLSGLAFRVFCHCTQYVRHSDDPLYTPETPTCHELLGHVPLLAEPFAQFS 290
QY 219 INMGLRFTKVIKQVQALPSKKORIQTQSNLIAIVRCFWFTVESGLIENHEGRKAYGAVL 278
DB 291 QEIG-----LASLGASEETVQKLAT-----CYFTVEFGCKQDGLRVFGAGL 334
QY 279 ISSPQELGHAFIDNVRVLPLELDQIRLPNTSTPQETLSIRHFDLVE 328
DB 335 LSSISELHLSGHAKVFPDPKACKQECILTSFDQVYFVSESFEDAKE 384

RESULT 8
TPH1 MOUSE
ID TPH1_MOUSE STANDARD; PRT; 447 AA.
AC P17532;
DT 01-AUG-1990 (Rel. 15, Created)
DT 01-AUG-1990 (Rel. 15, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Tryptophan 5-hydroxylase 1 (EC 1.14.16.4) (Tryptophan 5-
DE monooxygenase 1).
GN TPH1 OR TPH.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90243261; PubMed=2110547;
RA Stoll J., Kozak C.A., Goldman D.;
RT "Characterization and chromosomal mapping of a cDNA encoding
RT tryptophan hydroxylase from a mouse mastocytoma cell line.";
RL Genomics 7:88-96(1990).
CC -!- CATALYTIC ACTIVITY: L-tryptophan + 4-alpha-hydroxytetrahydrobiopterin.
CC -!- hydroxy-L-tryptophan + ferrous ion.
CC -!- COFACTOR: Ferrous ion.
CC -!- PATHWAY: Serotonin biosynthesis; first step.
CC -!- PATHWAY: Melatonin biosynthesis; first step.
CC -!- SUBUNIT: Multimer of identical subunits.
CC -!- SIMILARITY: Belongs to the bioprotein-dependent aromatic amino acid
CC hydroxylase family.

```

```

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; J04758; AAA63401.1; -.
CC PIR; A34582; A34582.
CC HSSP; P04177; ITOH.
CC MGD; MGI.98796; Tph1.
CC InterPro; IPR001273; Aaa_hydroxylase.
CC InterPro; IPR002912; ACT.
CC InterPro; IPR005963; Tyr_5_monox.
CC Pfam; PF01842; ACT; 1.
CC Pfam; PF00351; bioprotein H; 1.
CC PRINTS; PR00372; FWHYDRXLA.
CC ProDom; PD002559; Aaa_hydroxylase; 1.
CC TIGRFAMs; TIGR01270; Trp_5_monoox; 1.
CC PROSITE; PS00367; BIOPROTEIN HYDROXYL; 1.
CC Oxidoreductase; Monooxygenase; Serotonin biosynthesis; Iron;
KW Phosphorylation.
FT MOD_RES 61 PHOSPHORYLATION (BY PKA) (POTENTIAL).
FT METAL 275 IRON (BY SIMILARITY).
FT METAL 280 IRON (BY SIMILARITY).
FT METAL 320 IRON (BY SIMILARITY).
SQ SEQUENCE 447 AA; 51343 MW; 16C939F22A138BCA CRC64;

Query Match 12.0%; Score 226.5; DB 1; Length 447;
Best Local Similarity 27.4%; Pred. No. 2.1e-10;
Matches 63; Conservative 43; Mismatches 97; Indels 27; Gaps 4;

QY 107 WYRLSSRFLSKYSCPRFFLDYLEAFGLSLDFLDH-----QAVIKFFLETHFSYY 158
DB 177 WGTIFRELNKLPHYTHACR--EYLRLPLLSKYCGYREDNVPQLEDVSNFLKERTGFSIR 233
QY 159 PVSGFVAPHQVLSLQDRYFPIASVMTLDKNFSLTPDLIHLLGHVPLHLPFSSEFF 218
DB 234 PVAGYLSPRDFLSGLAFRVFCHCTQYVRHSDDPLYTPETPTCHELLGHVPLLAEPFAQFS 293
QY 219 INMGLRFTKVIKQVQALPSKKORIQTQSNLIAIVRCFWFTVESGLIENHEGRKAYGAVL 278
DB 294 QEIG-----LASLGASEETVQKLAT-----CYFTVEFGCKQDGLRVFGAGL 337
QY 279 ISSPQELGHAFIDNVRVLPLELDQIRLPNTSTPQETLSIRHFDLVE 328
DB 338 LSSISELHLSGHAKVFPDPKACKQECILTSFDQVYFVSESFEDAKE 387

RESULT 9
TPH1 RABIT
ID TPH1_RABIT STANDARD; PRT; 444 AA.
AC P17290; Q29523;
DT 01-AUG-1990 (Rel. 15, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Tryptophan 5-hydroxylase 1 (EC 1.14.16.4) (Tryptophan 5-
DE monooxygenase 1).
GN TPH1 OR TPH.
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_TaxID=9986;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=87289638; PubMed=3475690;
RA Grenett H.E., Ledley F.D., Reed L.L., Woo S.L.C.;
RT "Full-length cDNA for rabbit tryptophan hydroxylase: functional
RT domains and evolution of aromatic amino acid hydroxylases.";
RL Proc. Natl. Acad. Sci. U.S.A. 84:5530-5534(1987).
RN [2]
RP SEQUENCE FROM N.A.

```



```

RC TISSUE=Brain;
RX MEDLINE=95077422; PubMed=7986090;
RA Tipper J.P., Citron B.A., Ribeiro P., Kaufman S.;
RT "Cloning and expression of rabbit and human brain tryptophan
RL hydroxylase cDNA in Escherichia coli.";
RL Arch. Biochem. Biophys. 315:445-453(1994).
CC -|- CATALYTIC ACTIVITY: L-tryptophan + tetrahydrobiopterin + O(2) = 5-
CC hydroxy-L-tryptophan + 4-alpha-hydroxytetrahydrobiopterin.
CC -|- COFACTOR: Ferrous ion.
CC -|- PATHWAY: Serotonin biosynthesis; first step.
CC -|- PATHWAY: Melatonin biosynthesis; first step.
CC -|- SUBUNIT: Multimer of identical subunits.
CC -|- SIMILARITY: Belongs to the biotpterin-dependent aromatic amino acid
CC hydroxylase family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: M17250; AAA31487.1; -.
CC EMBL: L29305; AAA67051.1; -.
CC PIR: S51199; S51199.
CC HSSP: P04177; ITOH.
CC InterPro: IPR001273; Aaa_hydroxylase.
CC InterPro: IPR005963; Tyr_5_monox.
CC Pfam: PF01842; ACT; 1.
CC Pfam: PF00351; bioterin_H; 1.
CC PRINTS: PR00372; FWHYDRXLASE.
CC ProDom: PD002559; Aaa_hydroxylase; 1.
CC TIGRfam: TIGR01270; Tyr_5_monox; 1.
CC PROSITE: PS00367; BIOTPTERIN HYDROXYL; 1.
CC Oxidoreductase; Monooxygenase; Serotonin biosynthesis; Iron;
KW Phosphorylation.
KW MOD_RES 58 58 PHOSPHORYLATION (BY PKA) (POTENTIAL).
FT METAL 272 272 IRON (BY SIMILARITY).
FT METAL 277 277 IRON (BY SIMILARITY).
FT METAL 317 317 IRON (BY SIMILARITY).
FT METAL 102 102 M -> L (IN REF. 1).
FT CONFLICT 151 151 L -> S (IN REF. 2).
FT CONFLICT 202 203 KY -> ND (IN REF. 1).
FT CONFLICT 207 207 R -> Q (IN REF. 2).
FT CONFLICT 290 290 T -> K (IN REF. 1).
SQ SEQUENCE 444 AA; 51118 MW; Bf182451B28ECD80 CRC64;

Query Match 11.9%; Score 225.5; DB 1; Length 444;
Best Local Similarity 27.0%; Pred. No. 2.5e-10;
Matches 62; Conservative 44; Mismatches 97; Indels 27; Gaps 4;

QY 107 WYRLSSRFSLWKSVCPRFFLDYLEAFGLLSDFLDH-----QAVIKFFELETHSY 158
DB 174 WGTVFRELKLYPHTHAC---EYLKNUPLSKYCGYREDNIPQLEDISNFKERTGSIR 230
QY 159 PVSQFVAPHQYLSLQDRYFFIASVMRTLDKDNFSLTPDLIHDLLGHVPMLLHPSFSEFF 218
DB 231 PVAGYLSRDFSLGAFVHFCTQVVRHSSDPFYTPRPTDCHELLGHVPLLAESPAPQS 290
QY 219 INMGRFLTKVIEKQVALPSKKQRTQTLQSNLIAIVRCFWFTVESGLIENHGRKAYGAVL 278
DB 291 QEIG-----LASLGASEAVQKLT-----CYFFTVFGLCKQDGQURVFGAGL 334
QY 279 ISSPQELGHAFIDNVRLVPLDQIIRLPFNTSTPQETLFSIRHFDLVE 328
DB 335 LSSISELKHVLSGHAKVFPFPKITYKQECILITFDQVYFVSFSEFPAKE 384

RESULT 10
TY3H_ANGAN
ID TY3H_ANGAN STANDARD; PRT; 488 AA.

```

```

AC O42091;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DE 10-OCT-2003 (Rel. 42, Last annotation update)
DE Tyrosine 3-monooxygenase [EC 1.14.16.2] (Tyrosine 3-hydroxylase) (TH).
GN TH.
OS Anguilla anguilla (European freshwater eel).
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Anguilliformes; Anguillidae;
OC Anguilla.
OX NCBI_TaxID=7936;
RN [1]
RP SEQUENCE FROM N.A. PubMed=9681435;
RX MEDLINE=98344760; PubMed=9681435;
RA Boularand S., Biguet N.F., Vidal B., Veron M., Mallet J.,
RA Vincent J.D., Dufour S., Vernier P.;
RT "Tyrosine hydroxylase in the european eel (Anguilla anguilla): cDNA
RT cloning, brain distribution, and phylogenetic analysis.";
RL J. Neurochem. 71:460-470(1998).
CC -|- FUNCTION: plays an important role in the physiology of adrenergic
CC neurones (By similarity).
CC -|- CATALYTIC ACTIVITY: L-tyrosine + tetrahydropteridine + O(2) = 3,4-
CC dihydroxy-L-phenylalanine + dihydropteridine + H(2)O.
CC -|- COFACTOR: Ferrous ion.
CC -|- ENZYME REGULATION: Phosphorylation leads to an increase in the
CC catalytic activity (By similarity).
CC -|- PATHWAY: Catecholamine biosynthesis; first step.
CC -|- SIMILARITY: Belongs to the biotpterin-dependent aromatic amino acid
CC hydroxylase family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: AJ000731; CAA04264.1; -.
CC HSSP: P04177; ITOH.
CC InterPro: IPR001273; Aaa_hydroxylase.
CC InterPro: IPR005962; Tyr_3_monox.
CC Pfam: PF00351; bioterin_H; 1.
CC PRINTS: PR00372; FWHYDRXLASE.
CC ProDom: PD002559; Aaa_hydroxylase; 1.
CC TIGRfam: TIGR01269; Tyr_3_monox; 1.
CC PROSITE: PS00367; BIOTPTERIN HYDROXYL; 1.
CC Catecholamine biosynthesis; Oxidoreductase; Monooxygenase; Iron;
KW Neurotransmitter biosynthesis; Phosphorylation.
KW MOD_RES 38 38 PHOSPHORYLATION (BY PKA) (BY SIMILARITY).
FT METAL 321 321 IRON (BY SIMILARITY).
FT METAL 326 326 IRON (BY SIMILARITY).
FT METAL 366 366 IRON (BY SIMILARITY).
SQ SEQUENCE 488 AA; 55490 MW; 573553BA39EBD48 CRC64;

Query Match 11.9%; Score 225.5; DB 1; Length 488;
Best Local Similarity 28.1%; Pred. No. 2.8e-10;
Matches 65; Conservative 38; Mismatches 99; Indels 29; Gaps 5;

QY 107 WYRLSSRFSLWKSVCPRFFLDYLEAFGLLSDFLDH-----QAVIKFFELETHSY 157
DB 223 WREVSVTLRDLTYTHACS-----EHLERFLLERHCGYSNPQLQEDYSHFLKERTGQL 278
QY 158 YPVSGFVAPHQYLSLQDRYFFIASVMRTLDKDNFSLTPDLIHDLLGHVPMLLHPSFSEF 217
DB 279 RPVAGLLSARDFSLGAFVHFCTQVVRHSSDPFYTPRPTDCHELLGHVPLLAESPAPQS 338
QY 218 INMGRFLTKVIEKQVALPSKKQRTQTLQSNLIAIVRCFWFTVESGLIENHGRKAYGAV 277
DB 339 SQNIG-----LASLGASEEDIEKLT-----LYWFFVEGLCKQDGQVXKAYGAG 382
QY 278 LISSPQELGHAFIDNVRLVPLDQIIRLPFNTSTPQETLFSIRHFDLVE 328

```


383 LLSSYGELVHLSDEPERPEAAAAEPYQDNQYQSVFVSSFTDAKE 433

RESULT 11

```
TY3H MOUSE STANDARD; PRT; 498 AA.
AC P24529;
DT 01-MAR-1992 (Rel. 21, Created)
DT 01-MAR-1992 (Rel. 21, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Tyrosine 3-monooxygenase (EC 1.14.16.2) (Tyrosine 3-hydroxylase) (TH).
GN TH.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_TaxID=10090;
[1]
SEQUENCE FROM N.A.
RX MEDLINE=91248263; PubMed=1674869;
RA Ichikawa S., Sasacka T., Nagatsu T.;
RT "Primary structure of mouse tyrosine hydroxylase deduced from its
RT cDNA.";
RL Biochem. Biophys. Res. Commun. 176:1610-1616(1991).
[2]
SEQUENCE OF 1-30 FROM N.A.
RC STRAIN=BALB/c;
RA Morgan W.W., Bermudez J., Sharp Z.D.;
RL Submitted (APR-1992) to the EMBL/GenBank/DBJ databases.
CC 1- FUNCTION: Plays an important role in the physiology of adrenergic
CC neurones.
CC 1- CATALYTIC ACTIVITY: L-tyrosine + tetrahydropteridine + O(2) = 3,4-
CC dihydroxy-L-phenylalanine + dihydropteridine + H(2)O.
CC 1- COFACTOR: Ferrous ion.
CC 1- ENZYME REGULATION: Phosphorylation leads to an increase in the
CC catalytic activity.
CC 1- PATHWAY: Catecholamine biosynthesis; first step.
CC 1- SIMILARITY: Belongs to the biopterin-dependent aromatic amino acid
CC hydroxylase family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; M69200; AAA40434.1; -.
CC EMBL; X53503; CAA37580.1; -.
CC PIR; JN0068; JN0068.
CC HSSP; P04177; 1TOH.
CC MGD; MGI:98735; Th.
CC
CC GO; GO:0006585; P: dopamine biosynthesis from tyrosine; IDA.
CC GO; GO:0007507; P: heart development; IMP.
CC GO; GO:0008016; P: regulation of heart rate; IMP.
CC InterPro; IPR001273; Aaa_hydroxylase.
CC InterPro; IPR005962; Tyr_3_monox.
CC Pfam; PF00351; bioplerin_H_1.
CC PRINTS; PR00372; FWHYDRXLASE.
CC ProDom; PD002559; Aaa_hydroxylase; 1.
CC TIGRFAMs; TIGR01269; Tyr_3_monoox; 1.
CC PROSITE; PS00367; BIOPTERIN_HYDROXYL; 1.
CC Catecholamine biosynthesis; Oxidoreductase; Monooxygenase; Iron;
CC Neurotransmitter biosynthesis; Phosphorylation.
CC MOD_RES 19 19 PHOSPHORYLATION (BY CAMK2) (BY
CC SIMILARITY)
CC MOD_RES 31 31 PHOSPHORYLATION (BY SIMILARITY).
CC MOD_RES 40 40 PHOSPHORYLATION (BY PKA) (BY SIMILARITY).
CC DOMAIN 51 59 POLY-ALA.
CC METAL 331 331 IRON (BY SIMILARITY).
CC METAL 336 336 IRON (BY SIMILARITY).
CC METAL 376 376 IRON (BY SIMILARITY).
CC SEQUENCE 498 AA; 55992 MW; 627901/9664F6DC6 CRC64;
```

```
Query Match 11.9%; Score 225.5; DB 1; Length 498;
Best Local Similarity 27.5%; Pred No. 2.8e-10;
Matches 66; Conservative 42; Mismatches 101; Indels 31; Gaps 5;

QY 107 WYRLSSRFSLMKSCYPRFFLDYLAFLGLSLDFLH-----QAVIKFFLETHFSY 158
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
233 WKEVATLGLVATHACR---EHLEAFQLLERYCGYREDSPQLEDVSHFLKERTGFQUR 289
QY 159 PVSGFVAPHQYLSLQDRYFPIASVMRTLDKDNFSLTPDLIHDLGHVFWLLHPSPFSEFF 218
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
290 PVAGLSAADFLASLAFRVFQCTQYIRHASSPMHSPFPPCCHELLGHVFWMLADRTFAQFS 349
QY 219 INNGRLFTVKIEKVALPSKQRIQTLQSNLLIAIVRCFTVESGLIENHEGRKAYGAVL 278
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
350 QDIG-----LASLGASDEIEKLSLST-----VYWFTEFGGLCKNGELKAYGAGL 393
QY 279 ISSPQLGHAFIDNVAVLPLELDQIIRLPNTSTPQETLFSIRHF-----DELVELTSKLE 334
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
394 LSSYGBLLHSLSEEPVRAFPDPTAAVQYQDCTQYQYVVFVSEFSDAKDLRNVASRIQ 453

RESULT 12
TPH_XENLA STANDARD; PRT; 481 AA.
ID TPH_XENLA
AC Q92142;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Tryptophan 5-hydroxylase (EC 1.14.16.4) (Tryptophan 5-monooxygenase).
GN TPH.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
[1]
SEQUENCE FROM N.A.
RP TISSUE=Retina;
RX MEDLINE=94246419; PubMed=8189245;
RA Green C.B., Besharse J.C.;
RL "Tryptophan hydroxylase expression is regulated by a circadian clock
RL in Xenopus laevis retina.";
RL J. Neurochem. 62:2420-2428(1994).
CC 1- CATALYTIC ACTIVITY: L-tryptophan + tetrahydrobiopterin + O(2) = 5-
CC hydroxy-L-tryptophan + 4-alpha-hydroxytetrahydrobiopterin.
CC 1- COFACTOR: Ferrous ion.
CC 1- PATHWAY: Serotonin biosynthesis; first step.
CC 1- PATHWAY: Melatonin biosynthesis; first step.
CC 1- SUBUNIT: Multimer of identical subunits (By similarity).
CC 1- SIMILARITY: Belongs to the biopterin-dependent aromatic amino acid
CC hydroxylase family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; L20679; AAA21306.1; -.
CC PIR; I51567; I51567.
CC HSSP; P04177; 1TOH.
CC InterPro; IPR001273; Aaa_hydroxylase.
CC InterPro; IPR002912; ACT_
CC InterPro; IPR005963; Tyr_5_monox.
CC Pfam; PF01842; ACT; 1.
CC Pfam; PF00351; bioplerin_H_1.
CC PRINTS; PR00372; FWHYDRXLASE.
CC ProDom; PD002559; Aaa_hydroxylase; 1.
CC TIGRFAMs; TIGR01270; Ttp_5_monoox; 1.
CC PROSITE; PS00367; BIOPTERIN_HYDROXYL; 1.
```


MEDLINE=88183482; PubMed=2895648;
Abate C., Smith J.A., Jon I.H.,
"Characterization of the catalytic domain of bovine adrenal tyrosine
hydroxylase.";
Biochem. Biophys. Res. Commun. 151:1446-1453(1988).
[4].
SEQUENCE OF 1-27.
T18SUB=Adrenal medulla;
MEDLINE=88163736; PubMed=2894860;
Haavik J., Andersson K.K., Petersson L., Flatmark T.;
"Soluble tyrosine hydroxylase (tyrosine 3-monooxygenase) from bovine
adrenal medulla: large-scale purification and physicochemical
properties.";
Biochim. Biophys. Acta 953:142-156(1988).
[5].
[6].
[7].
[8].
[9].
[10].
[11].
[12].
[13].
[14].
[15].
[16].
[17].
[18].
[19].
[20].
[21].
[22].
[23].
[24].
[25].
[26].
[27].
[28].
[29].
[30].
[31].
[32].
[33].
[34].
[35].
[36].
[37].
[38].
[39].
[40].
[41].
[42].
[43].
[44].
[45].
[46].
[47].
[48].
[49].
[50].
[51].
[52].
[53].
[54].
[55].
[56].
[57].
[58].
[59].
[60].
[61].
[62].
[63].
[64].
[65].
[66].
[67].
[68].
[69].
[70].
[71].
[72].
[73].
[74].
[75].
[76].
[77].
[78].
[79].
[80].
[81].
[82].
[83].
[84].
[85].
[86].
[87].
[88].
[89].
[90].
[91].
[92].
[93].
[94].
[95].
[96].
[97].
[98].
[99].
[100].
[101].
[102].
[103].
[104].
[105].
[106].
[107].
[108].
[109].
[110].
[111].
[112].
[113].
[114].
[115].
[116].
[117].
[118].
[119].
[120].
[121].
[122].
[123].
[124].
[125].
[126].
[127].
[128].
[129].
[130].
[131].
[132].
[133].
[134].
[135].
[136].
[137].
[138].
[139].
[140].
[141].
[142].
[143].
[144].
[145].
[146].
[147].
[148].
[149].
[150].
[151].
[152].
[153].
[154].
[155].
[156].
[157].
[158].
[159].
[160].
[161].
[162].
[163].
[164].
[165].
[166].
[167].
[168].
[169].
[170].
[171].
[172].
[173].
[174].
[175].
[176].
[177].
[178].
[179].
[180].
[181].
[182].
[183].
[184].
[185].
[186].
[187].
[188].
[189].
[190].
[191].
[192].
[193].
[194].
[195].
[196].
[197].
[198].
[199].
[200].
[201].
[202].
[203].
[204].
[205].
[206].
[207].
[208].
[209].
[210].
[211].
[212].
[213].
[214].
[215].
[216].
[217].
[218].
[219].
[220].
[221].
[222].
[223].
[224].
[225].
[226].
[227].
[228].
[229].
[230].
[231].
[232].
[233].
[234].
[235].
[236].
[237].
[238].
[239].
[240].
[241].
[242].
[243].
[244].
[245].
[246].
[247].
[248].
[249].
[250].
[251].
[252].
[253].
[254].
[255].
[256].
[257].
[258].
[259].
[260].
[261].
[262].
[263].
[264].
[265].
[266].
[267].
[268].
[269].
[270].
[271].
[272].
[273].
[274].
[275].
[276].
[277].
[278].
[279].
[280].
[281].
[282].
[283].
[284].
[285].
[286].
[287].
[288].
[289].
[290].
[291].
[292].
[293].
[294].
[295].
[296].
[297].
[298].
[299].
[300].
[301].
[302].
[303].
[304].
[305].
[306].
[307].
[308].
[309].
[310].
[311].
[312].
[313].
[314].
[315].
[316].
[317].
[318].
[319].
[320].
[321].
[322].
[323].
[324].
[325].
[326].
[327].
[328].
[329].
[330].
[331].
[332].
[333].
[334].
[335].
[336].
[337].
[338].
[339].
[340].
[341].
[342].
[343].
[344].
[345].
[346].
[347].
[348].
[349].
[350].
[351].
[352].
[353].
[354].
[355].
[356].
[357].
[358].
[359].
[360].
[361].
[362].
[363].
[364].
[365].
[366].
[367].
[368].
[369].
[370].
[371].
[372].
[373].
[374].
[375].
[376].
[377].
[378].
[379].
[380].
[381].
[382].
[383].
[384].
[385].
[386].
[387].
[388].
[389].
[390].
[391].
[392].
[393].
[394].
[395].
[396].
[397].
[398].
[399].
[400].
[401].
[402].
[403].
[404].
[405].
[406].
[407].
[408].
[409].
[410].
[411].
[412].
[413].
[414].
[415].
[416].
[417].
[418].
[419].
[420].
[421].
[422].
[423].
[424].
[425].
[426].
[427].
[428].
[429].
[430].
[431].
[432].
[433].
[434].
[435].
[436].
[437].
[438].
[439].
[440].
[441].
[442].
[443].
[444].
[445].
[446].
[447].
[448].
[449].
[450].
[451].
[452].
[453].
[454].
[455].
[456].
[457].
[458].
[459].
[460].
[461].
[462].
[463].
[464].
[465].
[466].
[467].
[468].
[469].
[470].
[471].
[472].
[473].
[474].
[475].
[476].
[477].
[478].
[479].
[480].
[481].
[482].
[483].
[484].
[485].
[486].
[487].
[488].
[489].
[490].
[491].
[492].
[493].
[494].
[495].
[496].
[497].
[498].

Db282 PAAGLLSARDFLASLAFRVQCQYIRHASSPMHSPPECCHELLGHVEMLADRTCAQFS341

Qy219 INNGRLFTKVIKQALPKKKQRIQTLQSNLAIIVRCFWFTVESGLIENHGRKAYGAVL278

Db342 QDIG-----LASLGVSDSEIEKLST-----LYWFTVEFGLCKQNGEVKAYGAGL385

Qy279 ISSPQELGHAFIDNVRLVPLELQIIRLPNTSTPQETLFSIRHF-----DELVELTSKLE334

Db386 LSSYGELLSLSEPEIRAFDPDAAVQVPYQDTQYVVFVSESFSDAKDKLSYASRIQ445

RESULT 15

PH4H_RHILO

AC PH4H_RHILO STANDARD; PRT; 275 AA.

DT Q38D72;

DT 28-FEB-2003 (Rel. 41, Created)

DT 28-FEB-2003 (Rel. 41, Last sequence update)

DT 10-OCT-2003 (Rel. 42, Last annotation update)

DE Phenylalanine-4-hydroxylase (EC 1.14.16.1) (PAH) (Phe-4-monoxygenase).

GN PHA OR MLR4831.

OS Rhizobium loti (Mesorhizobium loti).

OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;

OC Phyllobacteriaceae; Mesorhizobium.

OX NCBI_TaxID=381;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=MAPPF303099;

RX MEDLINE=21082930; PubMed=11214968;

RA Kaneko T., Nakamura Y., Sato S., Asamizu E., Kato T., Sasamoto S.,

RA Watanabe A., Idegawa K., Ishikawa A., Kawashima K., Kimura T.,

RA Kishida Y., Kiyokawa C., Kohara M., Matsumoto M., Matsuno A.,

RA Mochizuki Y., Nakayama S., Nakazaki N., Shimpo S., Sugimoto M.,

RA Takeuchi C., Yamada M., Tabata S.;

RT "Complete genome structure of the nitrogen-fixing symbiotic bacterium

RT Mesorhizobium loti";

RL DNA Res. 7:331-338(2000).

CC -1- CATALYTIC ACTIVITY: L-phenylalanine + tetrahydrobiopterin + O(2) =

CC L-tyrosine + 4-alpha-hydroxytetrahydrobiopterin.

CC -1- COFACTOR: Binds 1 ferrous ion (by similarity).

CC -1- PATHWAY: Catabolism of phenylalanine; first (rate-limiting) step.

CC -1- SIMILARITY: Belongs to the biopterin-dependent aromatic amino acid

CC hydroxylase family.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to license@sib-sib.ch).

CC -----

DR EMBL; AF003005; BAB51399.1; -

DR InterPro; IPR001273; Aaa_hydroxylase.

DR InterPro; IPR005960; Phenylalaa4OaseW.

DR Pfam; PF00351; biopterin H; 1.

DR PRINTS; PR00372; FYWHYDRXLASE.

DR ProDom; PD002559; Aaa_hydroxylase; 1.

DR TIGRFAMs; TIGR01267; Phe4hydrox mono; 1.

DR PROSITE; PS00367; BIOPTERIN_HYDROXYL; 1.

DR Oxidoreductase; Monooxygenase; Phenylalanine catabolism; Iron;

KW Complete proteome.

FT METAL 135 IRON (POTENTIAL).

FT METAL 140 IRON (POTENTIAL).

SQ SEQUENCE 275 AA; 31347 MW; BC29D255534BC215 CRC64;

Qy139 FLDRQAVIKFFETHTFSPYVSGFVAPHQYLSLLQDRYFPPIASVMRTLDKDNFSLTDDL198

Db74 IPDEFEDYSTKLKLTGWEIIIAVPCGLIPAAFFFDHLANRRFPVTNWLRTQELDYIVEPDM133

Qy199 IHDLLGHVPLLLHPSFSEFFINMGRLETKVIEKVALPSKKQRIQTLQSNLIIAIVRCFWF258

Db134 PHDFFGHVPLSQVPVAFDM-----QMYGKKAGDIIALGGDEM-----ITRLYWI178

Qy259 TVESGLI-ENHEGRKAYGAVLISSPQELGHAFI-DNVRVLPFLDQIIRLPNTSTPQET316

Db179 TAEYGLVQERAGPLKAFGAGLMSSFTLEQFAVEGKDAHHVFPDLETVMRTGYEIDKFORA238

Qy317 LFSIRHFEDEL326

Db239 YFVLPSFDAL248

Search completed: March 25, 2004, 14:11:15

Job time : 18 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 25, 2004, 14:10:43 ; Search time 45 Seconds

(without alignments)
2538.169 Million cell updates/sec

US-09-438-185a-1047

Title:

Perfect score: 1889

Sequence: 1 VHYCERTLPKYILKALKL.....ESTPLYNQEKYLSGFEVLCC 362

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL 25:
1: sp_archaea:
2: sp_bacteria:
3: sp_fungi:
4: sp_human:
5: sp_invertebrate:
6: sp_mammal:
7: sp_mhc:
8: sp_organelle:
9: sp_phase:
10: sp_plant:
11: sp_rodent:
12: sp_virus:
13: sp_vertebrate:
14: sp_unclassified:
15: sp_virus:
16: sp_bacteriap:
17: sp_archaeap:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|--------|----------------------|
| 1 | 543.5 | 28.8 | 277 | Q822G6 | Q822G6 chlamydomophi |
| 2 | 243.5 | 12.9 | 471 | Q8A71 | Q8A71 brachydanio |
| 3 | 237.5 | 12.6 | 491 | Q9P40 | Q9P40 gallus gall |
| 4 | 231.5 | 12.3 | 446 | Q8K31 | Q8K31 mesocricetu |
| 5 | 230.5 | 12.2 | 264 | Q871N0 | Q871N0 vibrio para |
| 6 | 229 | 12.1 | 271 | Q8EGD8 | Q8EGD8 shewanella |
| 7 | 228.5 | 12.1 | 263 | Q8D6S0 | Q8D6S0 vibrio vuln |
| 8 | 226.5 | 12.0 | 262 | Q88EH3 | Q88EH3 pseudomonas |
| 9 | 226 | 12.0 | 265 | Q885L0 | Q885L0 pseudomonas |
| 10 | 224.5 | 11.9 | 497 | Q86370 | Q86370 schistosoma |
| 11 | 223.5 | 11.8 | 584 | Q81LM9 | Q81LM9 bacillus an |
| 12 | 220.5 | 11.7 | 584 | Q81B84 | Q81B84 bacillus ce |
| 13 | 219.5 | 11.6 | 296 | Q8PQ25 | Q8PQ25 xanthomonas |
| 14 | 215.5 | 11.4 | 453 | Q91WV1 | Q91WV1 mus musculu |
| 15 | 214 | 11.3 | 296 | Q8PE27 | Q8PE27 xanthomonas |
| 16 | 213.5 | 11.3 | 452 | Q8TEY0 | Q8TEY0 homo sapien |

| | | | | | | |
|----|-------|------|------|----|--------|---------------------|
| 17 | 206.5 | 10.9 | 244 | 2 | Q9AG78 | Q9AG78 streptomyce |
| 18 | 206.5 | 10.9 | 329 | 5 | Q817F1 | Q817F1 caenorhabdi |
| 19 | 206.5 | 10.9 | 522 | 5 | Q23438 | Q23438 caenorhabdi |
| 20 | 206.5 | 10.9 | 532 | 5 | Q9XZD1 | Q9XZD1 caenorhabdi |
| 21 | 205.5 | 10.9 | 447 | 5 | Q81901 | Q81901 aedes aegypt |
| 22 | 195.5 | 10.3 | 449 | 13 | Q7SYH6 | Q7SYH6 brachydanio |
| 23 | 181.5 | 9.6 | 457 | 5 | Q9XQ05 | Q9XQ05 caenorhabdi |
| 24 | 178.5 | 9.4 | 450 | 5 | Q96947 | Q96947 geodia cydo |
| 25 | 178.5 | 9.4 | 555 | 5 | Q9W0K2 | Q9W0K2 drosophila |
| 26 | 177 | 9.4 | 495 | 2 | Q9FDC3 | Q9FDC3 myxococcus |
| 27 | 164 | 8.7 | 163 | 4 | Q86Y20 | Q86Y20 homo sapien |
| 28 | 157.5 | 8.3 | 198 | 6 | Q8MJ22 | Q8MJ22 eryctolagus |
| 29 | 157 | 8.3 | 438 | 5 | Q7498 | Q7498 branchiost |
| 30 | 151 | 8.0 | 171 | 6 | Q8MIU1 | Q8MIU1 sus scrofa |
| 31 | 135.5 | 7.2 | 250 | 6 | Q9SLQ6 | Q9SLQ6 equus cabal |
| 32 | 123.5 | 6.5 | 115 | 4 | Q81ZE2 | Q81ZE2 homo sapien |
| 33 | 123 | 6.5 | 323 | 5 | Q86LZ5 | Q86LZ5 caenorhabdi |
| 34 | 112.5 | 6.0 | 1909 | 5 | Q86IF3 | Q86IF3 dictyosteli |
| 35 | 109.5 | 5.8 | 1253 | 10 | Q9SCZ3 | Q9SCZ3 arabidopsis |
| 36 | 108.5 | 5.7 | 1501 | 5 | Q9VMZ3 | Q9VMZ3 drosophila |
| 37 | 106 | 5.6 | 129 | 13 | Q42428 | Q42428 lates calca |
| 38 | 106 | 5.6 | 504 | 8 | Q9GF30 | Q9GF30 cardamine r |
| 39 | 106 | 5.6 | 504 | 8 | Q9GF31 | Q9GF31 cardamine p |
| 40 | 105.5 | 5.6 | 1128 | 12 | Q7T3L2 | Q7T3L2 cryptophleb |
| 41 | 104 | 5.5 | 513 | 8 | Q95F33 | Q95F33 muhlenbergi |
| 42 | 104 | 5.5 | 513 | 8 | Q95F29 | Q95F29 enneapogon |
| 43 | 103.5 | 5.5 | 513 | 8 | Q95F39 | Q95F39 monanthochl |
| 44 | 103.5 | 5.5 | 2064 | 5 | Q9VP19 | Q9VP19 drosophila |
| 45 | 103 | 5.5 | 504 | 8 | Q9GF41 | Q9GF41 aubrieta de |

ALIGNMENTS

RESULT 1

Q822G6 ID Q822G6 PRELIMINARY; PRT; 277 AA.
AC Q822G6;
DT 01-JUN-2003 (TREMBLrel. 24, Created)
DT 01-JUN-2003 (TREMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE Hypothetical protein.
GN CCA00716.
OS Chlamydomophila caviae.
OC Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydomophila.
OX NCBI_TaxID=83557;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=GPIC;
RX MEDLINE=22569155; PubMed=12682364;
RA Read T.D., Myers G.S.A., Brunham R.C., Nelson W.C., Paulsen I.T.,
RA Heidelberg J., Holtzapple E., Khouri H., Federova N.B., Carty H.A.,
RA Umayam L.A., Haft D.H., Peterson J., Beanan M.J., White O.,
RA Salzberg S.L., Hsia R.-C., McClarty G., Rank R.G., Bavoil P.M.,
RA Fraser C.M.;
RT "Genome sequence of Chlamydomophila caviae (Chlamydia psittaci GPIC):
RT examining the role of niche-specific genes in the evolution of the
RL Nucleic Acids Res. 31:2134-2147(2003).
DR EMBL; AS016996; AAP05458.1; ...
DR TIGR; CCA00716; ...
DR GO; GO:0005506; F1ron ion binding; IPA.
DR GO; GO:0004497; F1monooxygenase activity; IEA.
DR GO; GO:0009072; P1aromatic amino acid family metabolism; IEA.
DR InterPro; IPR001273; Aaa hydroxylase.
DR PRINTS; PR00372; FVWVDRXLASE.
DR ProDom; PD002559; Aaa hydroxylase; 1.
KW Hypothetical protein-Complete proteome.
SQ SEQUENCE 277 AA; 32109 MW; 0AA5E2EBA4398E8B CRC64;

Query Match 28.8%; Score 543.5; DB 16; Length 277;
Best Local Similarity 43.0%; Pred. No. 5.3e-39;
Matches 108; Conservative 47; Mismatches 95; Indels 1; Gaps 1;

| RESULT 3 | Q9PU40 | PRELIMINARY; | PRT; | 491 AA. |
|----------|---|---|------|---------|
| ID | Q9PU40 | | | |
| AC | Q9PU40; | | | |
| DT | 01-MAY-2000 | (TREMBLrel. 13, Created) | | |
| DT | 01-MAY-2000 | (TREMBLrel. 13, Last sequence update) | | |
| DT | 01-OCT-2003 | (TREMBLrel. 25, Last annotation update) | | |
| DE | Tyrosine hydroxylase. | | | |
| CS | Gallus gallus (Chicken) | | | |
| OC | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; | | | |
| OC | Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae; | | | |
| OC | Gallus. | | | |
| NCBI | TaxID=9031; | | | |

RP SEQUENCE FROM N.A.
RA Ernberger U.;
RP Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
RL [2]
RN
RP SEQUENCE FROM N.A.
RX MEDLINE=96076133; PubMed=7577670;
RA Ernberger U., Patzke H., Tissier-Seta J.P., Reh T., Goridis C.,
RA Richter H.;
RT "the expression of tyrosine hydroxylase and the transcription factors
RT cphox-2 and Cash-1: evidence for distinct inductive steps in the
RT differentiation of chick sympathetic precursor cells.";
RL Mech. Dev. 52:125-136(1995).
DR EMBL; AJ251387; CAB62388.1; -.
DB PDB; S40109; S40109

HSP7; P041111; IGH.
DR GO: 0003506; F:iron ion binding; IEA.
DR GO: 0004511; F:tyrosine 3-monooxygenase activity; IEA.
DR GO: 0004511; F:tyrosine 3-monooxygenase activity; IEA.
DR GO: 0003072; P:aromatic amino acid family metabolism; IEA.
DR GO: 0004243; P:catecholamine biosynthesis; IEA.
DR InterPro: IPR001273; Aaa hydroxylase.
DR InterPro: IPR005962; Tyr_3_monox.
DR Pfam: PF00351; biopterin_H; 1.
DR Pfam: PF00372; FWHYDRXLASE.
DR ProDom: PD02559; Aaa hydroxylase; 1.
DR TIGRFAMS: TIGR01269; Tyr 3_monox; 1.
DR PROSITE: PS00367; BIOPTERIN HYDROXYL; 1.
DR SOURCE 491 AA: 56016 MW: 6322F2D50E745930 CRC64;
DR

Db 283 PVAGLLSARDFLASTAARVFQCTQVIRHASSPMHSPEDCCHELLGHVEMLADKTFQAQS 341
Qy 219 INNGRLFTVKIEKQVALPSKKORIQTLOSLLIAIVRCSEFTVESGLIENHEGRKAYGAVL 276
Db 343 QDTG-----LASLGATDEIEKIAATL-----TWFTVEFGCTQNGIIVRAYGAGL 386
Qy 279 ISSPQELGHAFIDNVRVPLELDQIRLPFNFTSTPQETLFSIRHFE 325
Db 387 LSSVGEITSLSPDEVRVDFDDAAAVQYODONTQPVYFVSESFSD 433

RESULT 4
Q8K3R1

```

ID Q8K3R1 PRELIMINARY; PRT; 446 AA.
AC Q8K3R1;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Tryptophane hydroxylase.
OS Mesocricetus auratus (Golden hamster).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
OC Mesocricetus.
OX NCBI_TaxID=10036;
RN [1]
RP SEQUENCE FROM N.A.
RA Pisarchik A.V., Slominski A.;
RT "Hamster tryptophan hydroxylase mRNA, complete coding sequence.";
RL Submitted (Mar-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY034600; AAK59708.1; -.
DR GO; GO:0016597; F:amino acid binding; IEA.
DR GO; GO:0005506; F:iron ion binding; IEA.
DR GO; GO:0004510; F:tryptophan 5-monooxygenase activity; IEA.
DR GO; GO:0009072; P:aromatic amino acid family metabolism; IEA.
DR GO; GO:0008152; P:metabolism; IEA.
DR GO; GO:0042427; P:serotonin biosynthesis; IEA.
DR InterPro; IPR001273; Aaa_hydroxylase.
DR InterPro; IPR002912; ACT.
DR Pfam; PF01842; ACT; 1.
DR Pfam; PF00351; bioterin H; 1.
DR PRINTS; PR00372; FYHDXKLASE;
DR PRODOM; PD002559; Aaa_hydroxylase; 1.
DR TIGRFAMs; TIGR01270; TYP_5_monoox; 1.
DR PROSITE; PS00367; BIOTERIN_HYDROXYL; 1.
SQ SEQUENCE 446 AA; 51042 MW; 70E6AB813B65352 CRC64;

Query Match 12.3%; Score 231.5; DB 11; Length 446;
Best Local Similarity 24.9%; Pred. No. 1.2e-11;
Matches 75; Conservative 47; Mismatches 112; Indels 67; Gaps 6;

QY 63 KCISILEPKNLLFV-----HLJSLKXNQREG-----CST 92
Db 112 KKISLDPCANRLMYGSELADHPGKDNVYRRKRYFAELAMNXXKHGDDIPKIEPTGE 171
QY 93 DMNVSTPFFNRLMYRLSSFLSKSKYCRFFLDYLAEGFLSDFLDH-----QA 144
Db 172 EIKTWGTFEFLNKLYPHACR-----EYLRSLPLSKYCGVREINPOLED 218
QY 145 VIKFPELETHFSYYPVSGVAPHQVLSLLODRYFPIASVMRTLDKDNFSLTPDLIHLLG 204
Db 219 VSNFLKECTGFSIRPVAGYLSFDRFLSGLAFRVNCTQYVRHSSDPLYTPEDTCHELLG 278
QY 205 HVPWLLHPSFSEFFINMGRLETKVIEKQALPSKKQRIOTLQSLNLIIVRCFWFTVESGL 264
Db 279 HVPLLAEPSFAQFSQIG-----LASLGASEDTVQKLAT-----CYPTFVDFGL 322
QY 265 IENHEGRKAYGAVLTSPPQELGHAFIDNVRLPDELQIIRLPNTSPQPTLFSIRHFD 324
Db 323 CKQDQQLRVFGAGLSSISLKHLSGHAKVPDPKPVACKQEGCLITTFQDVYVSSFE 382
QY 325 E 325
Db 383 D 383

RESULT 5
Q87INO PRELIMINARY; PRT; 264 AA.
AC Q87INO;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Phenylalanine-4-hydroxylase.
GN VPA0576.
OS Vibrio parahaemolyticus.

ID Q8EGD8 PRELIMINARY; PRT; 271 AA.
AC Q8EGD8;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Phenylalanine-4-hydroxylase.
GN PHA OR S01666.
OS Shewanella oneidensis.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Alteromonadales;
OC Alteromonadales; Shewanella.
OX NCBI_TaxID=70863;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=MR-1.
RC MEDLINE=2237686; PubMed=12368813;
RA Heidelberg J.F., Paulsen I.T., Nelson K.E., Gaidos E.J., Nelson W.C.,
RA Read T.D., Eisen J.A., Seshadri R., Ward N., Methe B., Clayton R.A.,
RA Meyer T., Tsapin A., Scott J., Beanan M., Brinkac L., Daugherty S.,
RA DeBoy R.T., Dodson R.J., Durkin A.S., Haft D.H., Klonay J.F.,
RA Madupu R., Peterson J.D., Umayam L.A., White O., Wolf A.M.,
RA Vamathevan J., Weidman J., Imprim M., Lee K., Berry K., Lee C.,
RA Mueller J., Khouri H., Gill J., Utterback T.R., McDonald L.A.,
```

```

OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
OC Vibrionaceae; Vibrio.
OX NCBI_TaxID=670;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=RMD 2210633 / Serotype O3:K6;
RC MEDLINE=22508454; PubMed=12620739;
RA Makino K., Oshima K., Kurokawa K., Yokoyama K., Uda T., Tagomori K.,
RA Iijima Y., Najima M., Nakano M., Yamashita A., Kubota Y., Kimura S.,
RA Tasunaga T., Honda T., Shinagawa H., Hattori M., Iida T.;
RT "Genome sequence of Vibrio parahaemolyticus: a pathogenic mechanism
RT distinct from that of V. cholerae.";
RL Lancet 361:743-749 (2003).
DR EMBL; AP005085; BAC61919.1; -.
DR GO; GO:0005506; F:iron ion binding; IEA.
DR GO; GO:0004497; F:monooxygenase activity; IEA.
DR GO; GO:0009072; P:aromatic amino acid family metabolism; IEA.
DR InterPro; IPR001273; Aaa_hydroxylase.
DR Pfam; PF00351; bioterin H; 1.
DR PRINTS; PR00372; FYHDXKLASE.
DR PROSITE; PS00367; BIOTERIN_HYDROXYL; 1.
KW Complete proteome.
SQ SEQUENCE 264 AA; 30504 MW; 12503C53D379594B CRC64;

Query Match 12.2%; Score 230.5; DB 16; Length 264;
Best Local Similarity 25.2%; Pred. No. 7.2e-12;
Matches 64; Conservative 50; Mismatches 115; Indels 25; Gaps 5;

QY 106 LMYRLSSRFLSKSKYCRFFLDYLAEGFLSDFLDHQAVIKFFLE-THFSYYPVSGFV 164
Db 24 IWSDLVTRQMSVIKERACDAYLGLLELNLPODRVPOQPEINRVLMTGQVPEVPALI 83
QY 165 APHQVLSLLODRYFPIASVMRTLDKDNFSLTPDLIHLLGHVFWLLHPSFSEFFINMGR 224
Db 84 DEDRPFENLGNKRFVATLRTREDFVLQEPDFEIFGHCAMLTHPBFAAFTHEYGL 143
QY 225 FTKVIEKQALPSKKQRIOTLQSLNLIIVRCFWFTVESGLIENHEGRKAYGAVLTSPPQ 284
Db 144 GQATPQKAY-----LARYLWFVFEGLVREGTKYKYGGLIUSPGE 187
QY 285 LGHAFIDNVRLPDELQIIRLPNTSPQPTLFSIRHFDL-----VELTSKLEWMLD 338
Db 188 TIVALESESAIRESFDLQVLRTPYRIDIMQPKYVVDVDFSQLFQISQLNLKQADLAIE 247
QY 339 QGLLESIPLYNQE 352
Db 248 AGLLP--PLFEPKE 259

RESULT 6
Q8EGD8 PRELIMINARY; PRT; 271 AA.
AC Q8EGD8;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Phenylalanine-4-hydroxylase.
GN PHA OR S01666.
OS Shewanella oneidensis.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Alteromonadales;
OC Alteromonadales; Shewanella.
OX NCBI_TaxID=70863;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=MR-1.
RC MEDLINE=2237686; PubMed=12368813;
RA Heidelberg J.F., Paulsen I.T., Nelson K.E., Gaidos E.J., Nelson W.C.,
RA Read T.D., Eisen J.A., Seshadri R., Ward N., Methe B., Clayton R.A.,
RA Meyer T., Tsapin A., Scott J., Beanan M., Brinkac L., Daugherty S.,
RA DeBoy R.T., Dodson R.J., Durkin A.S., Haft D.H., Klonay J.F.,
RA Madupu R., Peterson J.D., Umayam L.A., White O., Wolf A.M.,
RA Vamathevan J., Weidman J., Imprim M., Lee K., Berry K., Lee C.,
RA Mueller J., Khouri H., Gill J., Utterback T.R., McDonald L.A.,
```

RAU Feldblyum T.V., Smith H.O., Venter J.C., Nealson K.H., Fraser C.M.;
RT "Genome sequence of the dissimilatory metal ion-reducing bacterium
Shewanella oneidensis";
RL Nat. Biotechnol. 20:1118-1123(2002).
DR ENBL; AEO15612; AAN54721.1; -;
DR TIGR; S01668; -;
DR GO; GO:000506; F:iron ion binding; IEA.
DR GO; GO:0004505; P:phenylalanine 4-monooxygenase activity; IEA.
DR GO; GO:0009072; P:aromatic amino acid family metabolism; IEA.
DR GO; GO:0006559; P:phenylalanine catabolism; IEA.
DR InterPro; IPR001273; Aaa hydroxylase.
DR Pfam; PF00351; biotpterin_H; 1.
DR PRINTS; PR00372; FYWHYDRXLAASE.
DR PRODOM; PD002559; Aaa hydroxylase; 1.
DR TIGRFAMs; TIGR01267; Phehydrox_mono; 1.
DR PROSITE; PS00367; BIOPTERIN_HYDROXYL; 1.
DR Complete proteome.
SQ SEQUENCE 271 AA; 31027 MW; 7FD47F4393DED742 CRC64;

Query Match 12.1%; Score 229; DB 16; Length 271;
Best Local Similarity 26.3%; Pred. No. 1e-11;
Matches 65; Conservative 49; Mismatches 113; Indels 20; Gaps 6;

QY 93 DMAVSTPFNNRLWYLLSSRFLMSKSYCPREFLDYLEAFGLSLDFLDHQAVI-KPFL 151
DB 18 DSGVIHYQBEHDMRWLYARQAVNLPGRACKYEQGLDALAMPKDRIPQLAEIDKVLMA 77
QY 152 ETHFSYYPVSGFVAPHOYLSLQDRYFPIASVMTLDKNFSLTPDLIHLGLHVPWLLH 211
DB 78 TTGWKTADVFPALLISFGFFELLANKFPFVATFIRKEEDYLOEPDFHEIFGHCPLLN 137
QY 212 PSFSEFFINMGRFLTKVIEKQVQALPKKKRIQTLQSNLIAIVRCFWFTVESGLIENHEOR 271
DB 138 PSFAHFHMYGQLGNA-----SKEDR-----VFLARLYWFTVEFGLLKPEGE 181
QY 272 KA-YGAVLISPOBELGHAFIDNV-RVLPLELDQIIRLPNTSTPQSLTFSIRHFDLVEL 329
DB 182 LCYGGGILSPGELTYAMESQVPERKFPDLQVLRTPYRIDIMQPIYVYVIEHIDVDEI 241
QY 330 TSKLEWM 336
DB 242 -AKNDIM 247

RESULT 7
ID Q8D6S0 PRELIMINARY; PRT; 263 AA.
AC Q8D6S0;
DT 01-MAR-2003 (TREMELrel. 23, Created)
DT 01-MAR-2003 (TREMELrel. 23, Last sequence update)
DT 01-OCT-2003 (TREMELrel. 25, Last annotation update)
DE Phenylalanine-4-hydroxylase.
GN VV20455.
OS Vibrio vulnificus.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
OC Vibrionaceae; Vibrrio.
OX NCBI_TaxID=672;
RN [1]
RP SEQUENCE FROM N.A.
RQ STRAIN=CMCP6;
RA Rhee J.H., Kim S.Y., Chung S.S., Kim J.J., Moon Y.H., Jeong H.,
RA Choy H.E.;
RT "Complete genome sequence of Vibrio vulnificus CMCP6";
RL Submitted (DSC-2002) to the ENBL/GenBank/DBJ databases.
DR ENBL; AEO16809; AAC07409.1; -;
DR GO; GO:000506; F:iron ion binding; IEA.
DR GO; GO:0004505; P:phenylalanine 4-monooxygenase activity; IEA.
DR GO; GO:0009072; P:aromatic amino acid family metabolism; IEA.
DR GO; GO:0006559; P:phenylalanine catabolism; IEA.
DR InterPro; IPR001273; Aaa hydroxylase.
DR InterPro; IPR005960; Phenylala4OHaseM.
DR Pfam; PF00351; biotpterin_H; 1.
SQ SEQUENCE 263 AA; 30222 MW; 2BC5A6F7FIA8F93 CRC64;

Query Match 12.1%; Score 228.5; DB 16; Length 263;
Best Local Similarity 25.6%; Pred. No. 1e-11;
Matches 65; Conservative 52; Mismatches 112; Indels 25; Gaps 7;

QY 106 LWYLLSSRFLMSKSYCPREFLDYLEAFGLSLDFLDHQAVIKFFELE-THFSYYPVSGFV 164
DB 24 IWQDLITRQLSMIQGRACSAIDGLALLDLFTDRVPQLPEINQVLAESTGWSVEVPALI 83
QY 165 APHOYLSLQDRYFPIASVMTLDKNFSLTPDLIHLGLHVPWLLHPSSEFFINMGR 224
DB 84 NDRFNNLLANKFPFVATFIRKEEDYLOEPDFHEIFGHCPMLINPDFAFTETVYGL 143
QY 225 FTKVIEKQVQALPKKKRIQTLQSNLIAIVRCFWFTVESGLIENHEGRKAYGAVLISPOE 284
DB 144 -----GEQATP--KQRAY-----LARYWFTVEFGLRBEGLRIYGGVSSPGE 187
QY 285 LGHAFID-NVRVLELDQIIRLPNTSTPQSLTFSIRHFDL-----VELTSLKLEWMLD 338
DB 188 TVYALEDERPERAKFDIQTVLRTPYRIDIMQPKFVLDINALFKSLDKIDLMKEVEFAMS 247
QY 339 QGLLESIPLYNQE 352
DB 248 AGLLP--PLPEKE 259

RESULT 8
Q88EH3
ID Q88EH3 PRELIMINARY; PRT; 262 AA.
AC Q88EH3;
DT 01-JUN-2003 (TREMELrel. 24, Created)
DT 01-JUN-2003 (TREMELrel. 24, Last sequence update)
DT 01-OCT-2003 (TREMELrel. 25, Last annotation update)
DE Phenylalanine-4-hydroxylase.
GN PHA OR PP4490.
OS Pseudomonas putida (strain KT2440).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Pseudomonadaceae; Pseudomonas.
OX NCBI_TaxID=160488;
RN [1]
RP SEQUENCE FROM N.A.
RQ MEDLINE=22423060; PubMed=12534463;
RA Nelson K.E., Weinel C., Paulsen I.T., Dodson R.J., Hilbert H.,
RA Martins dos Santos V.A.P., Fouts D.E., Gill S.R., Pop M., Holmes M.,
RA Brinkac L., Beanan M., DeBoy R.T., Daugherty S., Kolonay J.,
RA Madupu R., Nelson W., White O., Peterson J., Khouri H., Hance I.,
RA Chris Lee P., Holtzapple E., Scanlan D., Tran K., Moazzar A., H.,
RA Utterback T., Rizzo M., Lee K., Kosack D., Moestl D., Wedler H.,
RA Lauber J., Stjepandic B., Hoheisel J., Straetz M., Heim S.,
RA Kiewitz C., Eisen J., Timmis K.N., Duisterhoef A., Tuemmler B.,
RA Fraser C.M.;
RT "Complete genome sequence and comparative analysis of the
metabolically versatile Pseudomonas putida KT2440";
RL Environ. Microbiol. 4:799-808(2002).
DR ENBL; AEO16790; AAN70065.1; -;
DR TIGR; PP4490; -;
DR GO; GO:000506; F:iron ion binding; IEA.
DR GO; GO:0004497; P:monooxygenase activity; IEA.
DR GO; GO:0009072; P:aromatic amino acid family metabolism; IEA.
DR InterPro; IPR001273; Aaa hydroxylase.
DR Pfam; PF00351; biotpterin_H; 1.
DR PRINTS; PR00372; FYWHYDRXLAASE.
DR PROSITE; PS00367; BIOPTERIN_HYDROXYL; 1.
DR Complete proteome.
SQ SEQUENCE 262 AA; 30099 MW; F574FD1545D02258 CRC64;

Query Match 12.0%; Score 226.5; DB 16; Length 262;

191 GILLSSPREAVYSSSAPEQPDPLEAMRTPVRIDILQPLFVFLPDLKRLFLDLAGEDIM 244
 QY 335 WMLDQGL 341
 DB 241 AKVHEGM 247

RESULT 10
 O96370 PRELIMINARY; PRT; 497 AA.

AC O96370;
 DT 01-MAY-1999 (TrEMBLrel. 10, Created)
 DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 GN Trypophan hydroxylase (EC 1.14.16.4).
 DE TrpH
 OS Schistosoma mansoni (Blood fluke).
 OC Eukaryota; Metazoa; Platyhelminthes; Trematoda; Digenea; Strigeidida;
 OC Schistosomatidae; Schistosomatidae; Schistosoma.
 OX NCBI_Taxid=6183;
 RN [1]
 SEQUENCE FROM N.A.
 RP STRAIN=Puerto Rican;
 RC MEDLINE=99348305; PubMed=10419488;
 RA Hamdan F.F., Ribeiro P.;
 RT "Characterization of a stable form of tryptophan hydroxylase from the
 human parasite Schistosoma mansoni.";
 RL J. Biol. Chem. 274:21746-21754 (1999).
 DR ENBL; AF031034; AAD01923.1; -
 DR HSSP; P04177; ITOH.
 DR GO; GO:0003506; F:iron ion binding; IEA.
 DR GO; GO:0004510; F:tryptophan 5-monooxygenase activity; IEA.
 DR GO; GO:0005072; P:aromatic amino acid family metabolism; IEA.
 DR GO; GO:0042427; P:serotonin biosynthesis; IEA.
 DR InterPro; IPR001273; Aaa hydroxylase.
 DR InterPro; IPR003963; Tyr_5_monox.
 DR Pfam; PF00351; bioplerin H; 1.
 DR PRINTS; PR00372; FYWHYDRXLASE.
 DR PRODOM; PD002559; Aaa hydroxylase; 1.
 DR TIGRFAMS; TIGR01270; Tip 5 monoox; 1.
 DR PROSITE; PS00367; BIOPTERIN_HYDROXYL; 1.
 KW Oxidoreductase.
 SQ SEQUENCE 497 AA; 57598 MW; F8964E4B4B2C361D CRC64;

Query Match 11.9%; Score 224.5; DB 5; Length 497;
 Best Local Similarity 28.5%; Pred. No. 5.4e-11;
 Matches 68; Conservative 34; Mismatches 108; Indels 29; Gaps 5

QY 96 VYSTPFFNNLWYRLSSRPSLWK-SYCPRPFLDYEAFLGSLGFLDH-----QAVI 146
 DB 201 IVEYTEIEKTTWGRIVRELTLYKTSACHEF---QKNLGLQDKAGYNEFDLPQLQVVS 256

QY 147 KFELETHFSYYPVSGFVAPHQYLSLQDRYFPFIASVNRITLKDKNLSLTDPDLTHDLLGHV 206
 DB 257 DELKARTGFCLEPVGAYLSARDFLSGLAFRVFYCTQIRHQADFFYTPPEPDCHELLGHV 316

QY 207 PMLHPSPSEFFINMGRFLTFTKVIKQALPSKKORIQTLOSNIATVRCFWFTVESGLIE 266
 DB 317 PMLADPKFARFQETG-----LASLGSDEEIKLAT-----CYFFIIEGLCR 360

QY 267 NHEGRKAYGAVILSPQELGHAFIDNVRLPLELDQIIRLPFNVTSTQETLFIHFDE 325
 DB 361 QDNQLKAYGAGLLSVAELQHALSDKAVIKFPIPMKVINEECLVTTFQNGYFTSSPFD 419

RESULT 11
 Q81LM9 PRELIMINARY; PRT; 584 AA.

AC Q81LM9;
 DT 01-JUN-2003 (TrEMBLrel. 24, Created)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Phenylalanine-4-hydroxylase, putative.

BA4586.
OS *Bacillus anthracis* (strain Ames).
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; *Bacillus*.
OX NCBI_TaxID=198094;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22608414; PubMed=1271629;
RA Read T.D., Peterson S.N., Tourasee N., Baillie L.W., Paulsen I.T.,
RA Nelson K.B., Tettein H., Fouts D.E., Eisen J.A., Gill S.R.,
RA Holtzapple E.K., Oktat O.A., Helgason E., Ristone J., Wu M.,
RA Kolonay J.F., Beaman M.J., Dodson R.J., Brinkac L.M., Gwinn M.,
RA DeBoy R.T., Madpu R., Daugherty S.C., Durkin A.S., Haft D.H.,
RA Nelson W.C., Peterson J.C., Pop M., Khouri H.M., Radune D.,
RA Bertou J.L., Mahamoud Y., Jiang L., Hance I.R., Weidman J.F.,
RA Berry K.J., Plaut R.D., Wolf A.M., Watkins K.L., Nierman W.C.,
RA Hazen A., Cline R., Redmond C., Thwaite J.B., White O., Salzberg S.L.,
RA Thomason B., Friedlander A.M., Koehler T.M., Hanna P.C., Kolsto A.-B.,
RA Fraser C.M.;
RT "the genome sequence of *Bacillus anthracis* Ames and comparison to
RT closely related bacteria.";
RL Nature 423:81-86(2003).
DR EMBL; AE017038; AAP28292.1; -;
DR TIGR; BA4586; -;
DR GO; GO:0005506; F:iron ion binding; IEA.
DR GO; GO:0004937; F:monooxygenase activity; IEA.
DR GO; GO:0009072; P:aromatic amino acid family metabolism; IEA.
DR InterPro; IPR001273; Aaa_hydroxylase.
DR Pfam; PF00351; bioprotein H; 1.
DR ProDom; PD002559; Aaa_hydroxylase; 2.
DR Complete proteome.
KW SEQUENCE 584 AA; 64361 MW; BDB968E8A8EA02B8A CRC64;
Query Match 11.7%; Score 223.5; DB 16; Length 584;
Best Local Similarity 23.6%; Pred. No. 8e-11;
Matches 65; Conservative 51; Mismatches 115; Indels 45; Gaps 6;
QY 97 VSTPFF-----NRNLVRLSSRFLSKVSCPRFFLDYLEAFGLSDFLDHOAVIKFFE 150
DB 15 VSTQHYDQYTPVNHAVRYIMRQNHSLFKDVAHPAYVNGLOSSGINID-----AIPKVEE 69
QY 151 L-----ETHSYFVPSGVFAPHQYLSLQDRYFIASVMRTLDKDNFSLPDLIHLGLH 205
DB 70 MNECLASGCGAVTIDGLIPGVAFDFQGHGLLPATDIRKVENIEYTPAPDIVHEAAGH 129
QY 206 VPWLLHPSFSEFFINMGSLFTKVI-----VIEKVALPSKKQ-----RIQTQSNL 249
DB 130 APILLDPTAYKYVRFQGIQAKAFSTKEHDAFEAVRTLIVKESPTSTDEVTAAENNV 189
QY 250 I-----AIVRCFWFTVESGLIENHEGRKAYGAVLISSPOELGHAFTDNRVL 296
DB 190 IEKQNLVSGLSAEQISLFWMTVEYGLIGDIDNPKIYGAGLLSSVGSKHCLTDAVEKV 249
QY 297 PLELDQIIRLFPNTSTPQETLFSIRHPDELVE 332
DB 250 PFSIEACTSTTYDVTRMQPQLFVCKSPFEELTEALEK 285
RESULT 12
Q818B4 ID Q818B4 PRELIMINARY; PRT; 584 AA.
AC Q818B4;
DT 01-JUN-2003 (TREMBLrel. 24, Created)
DT 01-JUN-2003 (TREMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE Phenylalanine-4-hydroxylase (EC 1.14.16.1).
GN BC4352.
OS *Bacillus cereus* (strain ATCC 14579 / DSM 31).
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; *Bacillus*.
OX NCBI_TaxID=226900;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22608415; PubMed=12721630;
RA Ivanova N., Sorokin A., Anderson I., Galleron N., Candelson B.,

RA Kapatral V., Bhattacharya A., Reznik G., Mikhailova N., Lapidus A.,
RA Chu L., Mazur M., Goltzman E., Larsen N., D'Souza M., Walunas T.,
RA Grechkin Y., Busch G., Haselkorn R., Fonstein M., Ehrlich S.D.,
RA Overbeek R., Kyrpides N.;
RT "Genome sequence of *Bacillus cereus* and comparative analysis with
RT *Bacillus anthracis*.";
RL Nature 423:87-91(2003).
DR EMBL; AE017012; AAP11265.1; -;
DR GO; GO:0005506; F:iron ion binding; IEA.
DR GO; GO:0004505; P:phenylalanine 4-monooxygenase activity; IEA.
DR GO; GO:0009072; P:aromatic amino acid family metabolism; IEA.
DR InterPro; IPR001273; Aaa_hydroxylase.
DR Pfam; PF00351; bioprotein H; 1.
DR ProDom; PD002559; Aaa_hydroxylase; 2.
DR Oxidoreductase; Complete proteome.
KW SEQUENCE 584 AA; 64225 MW; 561092C0323DFE89 CRC64;
Query Match 11.7%; Score 220.5; DB 16; Length 584;
Best Local Similarity 23.5%; Pred. No. 1.5e-10;
Matches 64; Conservative 51; Mismatches 112; Indels 45; Gaps 5;
QY 97 VSTPFF-----NRNLVRLSSRFLSKVSCPRFFLDYLEAFGLSDFLDHOAVIKFFE 150
DB 15 VSTQHYDQYTPVNHAVRYIMRQNHSLFKDVAHPAYVNGLOSSGINID-----AIPKVEE 69
QY 151 LB-----THFSYFVPSGVFAPHQYLSLQDRYFIASVMRTLDKDNFSLPDLIHLGLH 205
DB 70 MNECLASGCGAVTIDGLIPGVAFDFQGHGLLPATDIRKVENIEYTPAPDIVHEAAGH 129
QY 206 VPWLLHPSFSEFFINMGSLFTKVI-----EKKVALPS-- 237
DB 130 APILLDPTAYKYVRFQGIQAKAFSTKEHDAFEAVRTLIVKESPTSTDEVTAAENAV 189
QY 238 -KKQIRQLTQSNLIAIVRCFWFTVESGLIENHEGRKAYGAVLISSPOELGHAFTDNRVL 296
DB 190 IEKQNLVSGLSAEQISLFWMTVEYGLIGDIDNPKIYGAGLLSSVGSKHCLTDAVEKV 249
QY 297 PLELDQIIRLFPNTSTPQETLFSIRHPDELVE 328
DB 250 PFSIEACTSTTYDVTRMQPQLFVCKSPFEELID 281
RESULT 13
Q8PQZ5 ID Q8PQZ5 PRELIMINARY; PRT; 296 AA.
AC Q8PQZ5;
DT 01-OCT-2002 (TREMBLrel. 22, Created)
DT 01-OCT-2002 (TREMBLrel. 22, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE Phenylalanine hydroxylase.
GN PAH OR XAC0174.
OS *Xanthomonas axonopodis* (pv. citri).
OC Bacteria; Proteobacteria; Gammaproteobacteria; *Xanthomonadales*.
OX NCBI_TaxID=92829;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=306 / ATCC 13902 / XV 101;
RX MEDLINE=2202145; PubMed=12024217;
RA da Silva A.C.R., Ferro J.A., Reinach F.C., Farah C.S., Furlan L.R.,
RA Quaggio R.B., Monteiro-Vitorello C.B., Van Sluys M.A., Almeida N.F.,
RA Alves L.M.C., do Amaral A.M., Bertolini M.C., Camargo L.E.A.,
RA Camarotte G., Cannavan F., Cardozo J., Chambergo F., Ciapina L.P.,
RA Cicarelli R.M.B., Coutinho L.L., Cursino-Santos J.R., El-Dorry H.,
RA Faria J.B., Ferreira A.J.S., Ferreira R.C.C., Ferro M.I.T.,
RA Formighieri E.F., Franco M.C., Greggio C.C., Gruber A.,
RA Katsuyama A.M., Kishi L.T., Leite R.P., Lemos E.G.M., Lenos M.V.F.,
RA Locati E.C., Machado M.A., Madeira A.M.B.N., Martinez-Rossi N.M.,
RA Martins E.C., Meidanis J., Menck C.F.M., Miyaki C.Y., Moon D.H.,
RA Moreira L.M., Novo M.T.M., Okura J.A.D., Oliveira M.C., Oliveira V.R.,
RA Pereira H.A., Rossi A., Sena J.A.D., Silva C., de Souza R.F.,
RA Spinola L.A.F., Takita M.A., Tamura R.E., Teixeira E.C., Tezra R.I.D.,
RA Trindade dos Santos M., Truffi D., Tsai S.M., White F.F.,

```
RA Setubal J.C., Kitajima J.P.;
RT Comparison of the genomes of two Xanthomonas pathogens with differing
RL host specificities.
RL Nature 417:459-463(2002).
DR EMBL; AE011641; AAM35066.1; -.
DR GO; GO:0005506; F:iron ion binding; IEA.
DR GO; GO:0004505; P:phenylalanine 4-monooxygenase activity; IEA.
DR GO; GO:0009072; P:aromatic amino acid family metabolism; IEA.
DR GO; GO:0006559; P:phenylalanine catabolism; IEA.
DR InterPro; IPR001273; Aaa_hydroxylase.
DR InterPro; IPR005960; Phenylalanyl 4-OHaseM.
DR Pfam; PF00351; biotpterln H; 1.
DR PRINTS; PR00372; FYWYDRKLASE.
DR ProDom; PD002559; Aaa_hydroxylase; 1.
DR TIGRFAMs; TIGR01267; Phe4hydrox_mono; 1.
DR PROSITE; PS00367; BIOPTERIN_HYDROXYL; 1.
KW Complete proteome.
SQ SEQUENCE 296 AA; 33467 MW; 258B1955670BC4BD CRC64;

Query Match 11.6%; Score 219.5; DB 16; Length 296;
Best Local Similarity 27.9%; Pred. No. 7.5e-11;
Matches 61; Conservative 39; Mismatches 82; Indels 37; Gaps 5;

QY 126 FLYDLFAFLSLDFLDH-QAVIKFFELETHFSYPSGFFVAPHQYLSLLQDXYEPIASVM 184
DB 60 FLQAQDAMGNGQTHIPRFDALNRVLQAATGWTIVGVQGLLPDLDFDHLNRRPPTVWMI 119
QY 185 RTLDKDNFSLTPDLIHLGLHVPWLLHPSFSEFFINNRLFTKVIKQVQLPSKKQRIQT 244
DB 120 RRPQDIYIAEPDLFDLFGHVPLLNPLFADFPMQAYGR-----GGVKAHGIGPDALQN 173
QY 245 LQSNLIAVCFWFTVBSGJENHEGRKAYGAVLISS-----PQELGHAFIDN 292
DB 174 L-----TRLVYWTVEFGILDTPOGLRIYAGIVSGESLVSLESPAPNRIG----- 220
QY 293 VRVLPLELDQIIRLPFNTSTPQETLFSIRHFDDELVELTS 331
DB 221 -----FDLQRMTRYRIDSFQKTYFVIDSFAQLMEATA 254

RESULT 14
Q91WV1 PRELIMINARY; PRT; 453 AA.
AC Q91WV1
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Phenylalanine hydroxylase.
GN PAH.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
[1]
RN SEQUENCE FROM N.A.
RP Strausberg R.;
RC TISSUE=Kidney;
RL Submitted (SSP-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC013458; AAH13458.1; -.
DR MGD; MGI:97473; Pah.
DR GO; GO:0016597; F:amino acid binding; IEA.
DR GO; GO:0005506; F:iron ion binding; IEA.
DR GO; GO:0004505; P:phenylalanine 4-monooxygenase activity; IEA.
DR GO; GO:0009072; P:aromatic amino acid family metabolism; IEA.
DR GO; GO:0008152; P:metabolism; IEA.
DR GO; GO:0006559; P:phenylalanine catabolism; IEA.
DR InterPro; IPR001273; Aaa_hydroxylase.
DR InterPro; IPR002912; ACT.
DR InterPro; IPR005961; Phenylalanyl 4-OHaseM.
DR Pfam; PF01842; ACT; 1.
DR Pfam; PF00351; biotpterln H; 1.
DR PRINTS; PR00372; FYWYDRKLASE.
DR ProDom; PD002559; Aaa_hydroxylase; 1.

DR TIGRFAMs; TIGR01268; Phe4hydrox_tetr; 1.
DR PROSITE; PS00367; BIOPTERIN_HYDROXYL; 1.
DR SEQUENCE 453 AA; 51899 MW; 551F181FA59DEA5B CRC64;

Query Match 11.4%; Score 215.5; DB 11; Length 453;
Best Local Similarity 27.8%; Pred. No. 2.9e-10;
Matches 64; Conservative 40; Mismatches 105; Indels 21; Gaps 5;

QY 104 RNLM---YRLSSRPSFLWKSYCPRRFFLDYLEAF-GLLSDFLDH-QAVIKFFELETHFSY 158
DB 184 RKTGWTFTLKALYKTHACAYEHNNHIFPLEKCYGFPREDNIPQLEDYSQPLQCTCTGFLR 243
QY 159 PVSQFVAPHQYLSLQDRYFPFIASVMRTLDKDNFSLTPDLIHLGLHVPWLLHPSFSEFF 218
DB 244 PVAGLLSSRDFLGLAFKVFHCTQYIRHGSKPMYTPEDFICHELLGHVPLFSDRSFQAQS 303
QY 219 INMGRFLTQVIEKQVALPSKKQRIQTLSQNLIAVRCFWFTVBSGJENHEGRKAYGAVL 278
DB 304 QEIG-----LASLGAPDEYIEKLTAT-----IYWFVTEFLCKEGDSIKAYGAGL 347
QY 279 ISSFQELGHAFIDNVRVLPLELDQIIRLPFNTSTPQETLFSIRHFDDELVE 328
DB 348 LSSFCQELQCLSDKPKLPLELEKTAQCEYTVTFQPLYYVAESFNDAKE 397

RESULT 15
Q9PE27 PRELIMINARY; PRT; 296 AA.
AC Q9PE27
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Phenylalanine hydroxylase.
GN PAH OR XCC0156.
OS Xanthomonas campestris (pv. campestris).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
OC Xanthomonadaceae; Xanthomonas.
OX NCBI_TaxID=340;
[1]
RN SEQUENCE FROM N.A.
RP STRAIN=ATCC 33913 / NCPPB 528;
RX MEDLINE=2202145; PubMed=12024217;
RA da Silva A.C.R., Ferro J.A., Reinach F.C., Farah C.S., Furlan L.R.,
RA Quaggio R.B., Monteiro-Vitorello C.B., Van Sluys M.A., Almeida N.F.,
RA Alves L.M.C., do Amaral A.M., Bertolini M.C., Camargo L.B.A.,
RA Camarotte G., Cannavan F., Cardoso J., Chambergo F., Ciapina L.P.,
RA Cicarelli R.M.B., Coutinho L.L., Cursino-Santos J.R., El-Dorry H.,
RA Faria J.B., Ferreira A.J.S., Ferreira R.C.C., Ferro M.T.T.,
RA Formighieri E.F., Franco M.C., Greggio C.C., Gruber A.,
RA Katsuyama A.M., Kishi L.T., Leite R.P., Lemos E.G.M., Lemos M.V.F.,
RA Locati E.C., Machado M.A., Madeira A.M.B.N., Martinez-Rossi N.M.,
RA Martins E.C., Meidanis J., Menck C.F.M., Miyaki C.Y., Moon D.H.,
RA Moreira L.M., Novo W.T.M., Okura V.K., Oliveira M.C., Oliveira V.R.,
RA Pereira H.A., Rossi A., Sena J.A.D., Silva C., de Souza R.F.,
RA Spindola L.A.F., Takita M.A., Tamura R.E., Teixeira E.C., Tezra R.I.D.,
RA Trindade dos Santos M., Truffi D., Tsai S.M., White F.F.,
RA Setubal J.C., Kitajima J.P.;
RT "Comparison of the genomes of two Xanthomonas pathogens with differing
RL host specificities."
RL Nature 417:459-463(2002).
DR EMBL; AE011641; AAM39475.1; -.
DR GO; GO:0005506; F:iron ion binding; IEA.
DR GO; GO:0004505; P:phenylalanine 4-monooxygenase activity; IEA.
DR GO; GO:0009072; P:aromatic amino acid family metabolism; IEA.
DR GO; GO:0006559; P:phenylalanine catabolism; IEA.
DR InterPro; IPR001273; Aaa_hydroxylase.
DR InterPro; IPR005960; Phenylalanyl 4-OHaseM.
DR Pfam; PF00351; biotpterln H; 1.
DR PRINTS; PR00372; FYWYDRKLASE.
DR ProDom; PD002559; Aaa_hydroxylase; 1.
DR TIGRFAMs; TIGR01267; Phe4hydrox_mono; 1.
DR PROSITE; PS00367; BIOPTERIN_HYDROXYL; 1.
KW Complete proteome.
```

```

SQ SEQUENCE 296 AA; 33232 MW; C2A4PE38B24B33DF CRC64;
Query Match 11.3%; Score 214; DB 16; Length 296;
Best Local Similarity 26.9%; Pred. NO. 2.3e-10;
Matches 66; Conservative 4; Mismatches 97; Indels 38; Gaps 7;
Qy 126 FLDYLEAFGLSDFLDHQAVIKF-----FELETHFSYYPVSGFVAPHQVLSLLQDRYFP 179
Db 60 FLQAQDAMG-----MDDTQIPRFDALNAVLAQTGWTLVGVEGLLPDLDFDHLANRFP 114
Qy 180 IASVWRILDKNPSLTDDLHDLGHVPLHLSFSEFFINMGRLFTKVIKQVALPSKK 239
Db 115 VTWIRPDQIDYIAEPDLFHLFGHVPLLMNPLFADFMOAYGR-----GGVKAHGIGP 168
Qy 240 QRIQTLOSNIATVRCFWFTVESGLIENHEGRKAYGAVLISSPOELCHAFIDNV-EVLPL 298
Db 169 DALQNL-----TRLWYIVFGLIATPQGLRIYGAGIVSSKGSLSLESAPNRVGF 221
Qy 299 ELQIIRLPNTSTPQSTLFSIRHDELVELTSLKLEWMLDQGLLESIPLY---NQEKYLS 355
Db 222 DLQVMTTRYRIDSFQXTYFVIDSFTQLMDATAP-----DFTPIYAALAQQPVP 271
Qy 356 GFEVL 360
Db 272 AGEVL 276

```

Search completed: March 25, 2004, 14:13:25
Job time : 47 secs